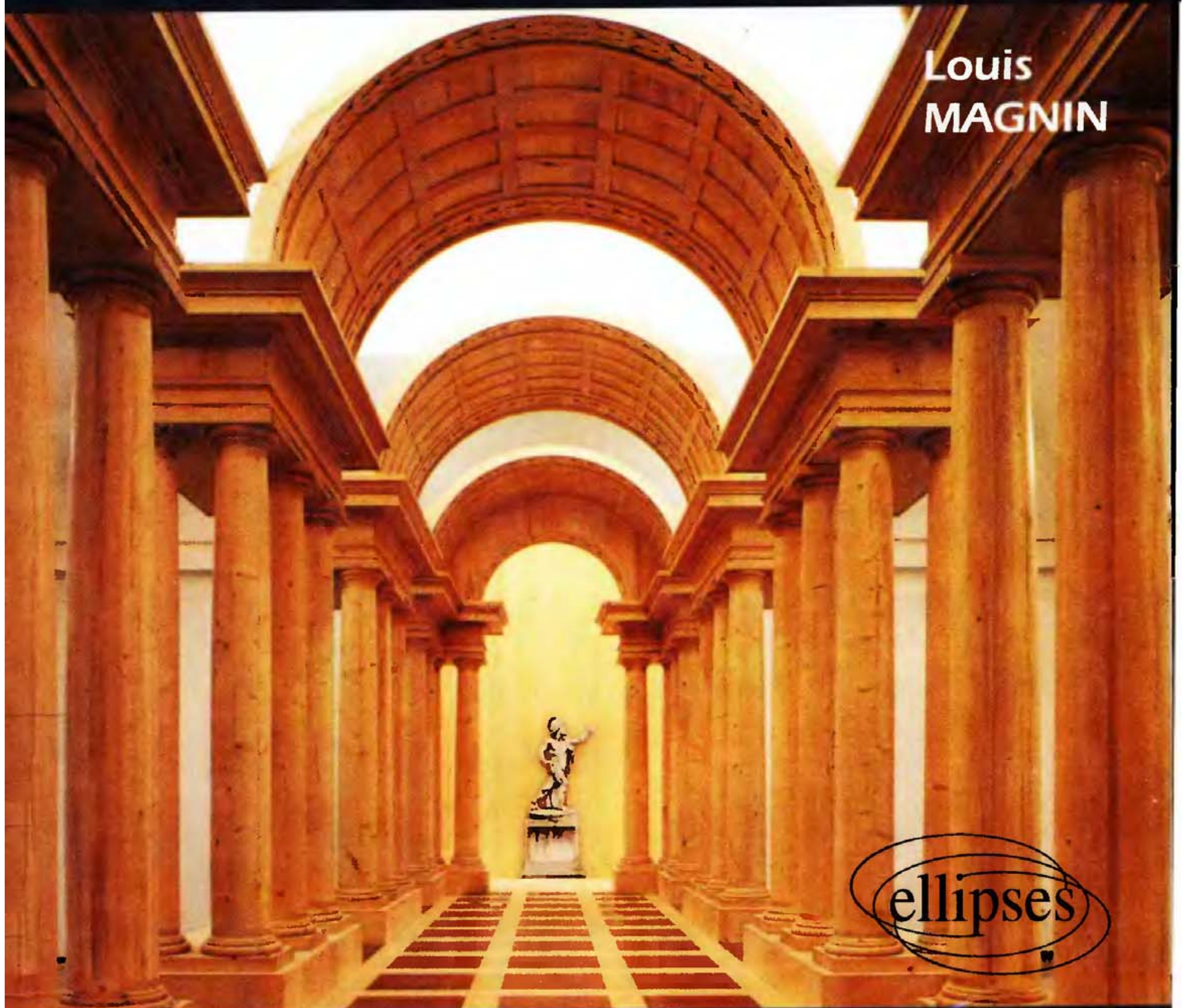
# Quelques questions d'Algèbre, Géométrie et Probabilités



# *Quelques questions* d'Algèbre, Géométrie et Probabilités

Louis MAGNIN



# Avant-propos

Ce livre a été élaboré à partir d'un cours enseigné depuis 1997 dans l'un des modules de mathématiques de la Licence Pluridisciplinaire de Sciences et Technologie de l'Université de Bourgogne.

Il s'adresse à des étudiants de second cycle, qui ne voudraient pas suivre un cycle spécialisé en mathématiques, mais désireraient acquérir une formation générale en mathématiques sur les sujets abordés ici, afin de pouvoir préparer certains concours. Il s'agit des concours de la catégorie A ou de concours de recrutement d'enseignants tels que CERPE (concours externe de recrutement des Professeurs des Ecoles) ou CAPLP2 (Certificat d'aptitude au Professorat des lycées professionnels).

Les prérequis sont très limités, et cet ouvrage est accessible aux étudiants ayant suivi un DEUG, un DUT ou même un BTS. Pour autant, il ne néglige pas la rigueur mathématique tout en restant dans le cadre fixé: les notions mathématiques mises en oeuvre sont développées intégralement, à l'exception de la théorie de l'intégrale de Lebesgue et de la démonstration de certains théorèmes limites qui dépassent le cadre fixé, à la fois par les notions à introduire et par la longueur nécessitée.

Les questions abordées s'articulent autour des thèmes suivants :

- Algèbre (groupes, actions de groupe, groupes de transformations),
- Géométrie (espace affine, coniques, mouvement newtonien, angles et cocyclicité),
- Probabilités (lois discrètes ou continues, théorèmes limites).

Chaque chapitre est suivi d'exercices, dont la plupart avec des indications formant un véritable corrigé.

Outre les étudiants mentionnés ci-dessus, cet ouvrage est susceptible d'intéresser un public assez large, allant du Grand Public souhaitant prendre contact avec des mathématiques avancées, jusqu'aux étudiants de Licence de mathématiques pures, qui pourront y retrouver certains points, tels que sections coniques, mouvement des planètes, mesure des angles, données explicites pour la comparaison de diverses approximations de la loi binômiale.

L'ouvrage est complété par des tables de lois de probabilité usuelles. Ces tables, ainsi que les autres données numériques concernant les Probabilités, ont été calculées avec le logiciel *REDUCE*.

La fin d'une démonstration est signalée par un □.

L.M

# Table des matières

# Chapitre 1

# Groupes, anneaux.

## 1.1 Définition d'un groupe.

**Définition 1. 1.** *On appelle groupe un ensemble $G$ muni d'une loi de composition interne*

$$\begin{array}{rcl} * : G \times G & \to & G \\ (x, y) & \mapsto & x * y \end{array}$$

*ayant les 3 propriétés suivantes.*
*(**A**) La loi $*$ est **associative** :*

$$x * (y * z) = (x * y) * z \quad \forall\, x, y, z \in G.$$

*(**N**) La loi $*$ possède un élément **neutre** :*

$$\exists e \in G \quad x * e = e * x \quad \forall\, x \in G.$$

*(**S**) Tout élément de $G$ possède un **symétrique** pour la loi $*$ :*

$$\forall x \in G \quad \exists x' \in G \quad x * x' = x' * x = e.$$

*Le groupe $G$ est dit commutatif (ou abélien) si en plus :*
*(**C**) La loi $*$ est **commutative** :*

$$x * y = y * x \quad \forall\, x, y \in G.$$

*Si $G$ est un groupe fini, son cardinal noté card $G$ ou $|G|$ sera encore appelé l'ordre de $G$.*

**Remarques.** (i) L'élément neutre $e$ d'un groupe $G$ est unique. En effet, si $e' \in G$ était un autre élément neutre, on aurait $e = e * e' = e'$.
(ii) Le symétrique d'un élément $x$ d'un groupe $G$ est unique. En effet, si $x', x'' \in G$ étaient deux éléments de $G$ symétriques de $x$, on aurait

$$x' = x' * e = x' * (x * x'') = (x' * x) * x'' = e * x'' = x''.$$

Si la loi $*$ est notée *multiplicativement* (*resp. additivement*), le symétrique d'un élément $x$ est appelé *inverse* (*resp. opposé*) et noté $x^{-1}$ (*resp.* $-x$). Le groupe $G$ est

alors dit *multiplicatif* (*resp. additif*). Si $G$ est additif, l'élément neutre est noté 0.
(iii) Pour $a \in G$, les *translations à gauche* (*resp. translations à droite*)

$$\begin{array}{rcl} L_a \,(resp.\ R_a) : G & \to & G \\ a & \mapsto & L_a(x) = a * x \,(resp.\ R_a(x) = x * a) \end{array}$$

sont des bijections de $G$ sur $G$. On a en effet pour tout $y \in G$ :

$$y = L_a(x) \Leftrightarrow y = a * x \Leftrightarrow a' * y = x$$

$$y = R_a(x) \Leftrightarrow y = x * a \Leftrightarrow y * a' = x$$

où $a'$ est le symétrique de $a$. En particulier, chacune des équations $a * x = e$ ou $x * a = e$ implique $x = a'$.
(iv) Pour $a, b \in G$, le symétrique de $a * b$ est $(a * b)' = b' * a'$. En effet,

$$(a * b) * (b' * a') = ((a * b) * b') * a' = (a * (b * b')) * a' = (a * e) * a' = a * a' = e.$$

(v) Un groupe possède au moins un élément, à savoir son élément neutre $e$.

## 1.2 Exemples.

(i) $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ munis de l'addition $+$ sont des groupes.

(ii) $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$, $\mathbb{R}^*_+ = ]0, +\infty[$, $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$, $\mathbb{T} = \{z \in \mathbb{C}; |z| = 1\}$, $\mu_n = \{z \in \mathbb{C};\ z^n = 1\}$ $(n \in \mathbb{N}^* = \mathbb{N} \setminus \{0\})$ munis de la multiplication $\cdot$ sont des groupes.

(iii) L'ensemble $M_n(\mathbb{C})$ (*resp.* $M_n(\mathbb{R}), M_n(\mathbb{Q}), M_n(\mathbb{Z})$) des matrices $n \times n$ à coefficients complexes (*resp.* réels, rationnels, entiers) muni de l'addition matricielle est un groupe.

(iv) L'ensemble $GL(n, \mathbb{C})$ (*resp.* $GL(n, \mathbb{R}), GL(n, \mathbb{Q})$) des matrices $n \times n$ *inversibles* à coefficients complexes (*resp.* réels, rationnels) muni de la multiplication matricielle est un groupe.

(v) L'ensemble $GL(n, \mathbb{Z})$ des matrices $n \times n$ *inversibles* à coefficients entiers, et dont les coefficients de la matrice inverse sont aussi à coefficients entiers, muni de la multiplication matricielle est un groupe.

## 1.3 Règles sur les puissances.

### 1.3.1 Lemme sur l'associativité.

Soit $G$ un ensemble muni d'une loi interne associative $*$. Pour $x, y, z \in G$, on pose

$$x * y * z = x * (y * z) = (x * y) * z.$$

De même, pour tout $n \geqslant 3$ on définit par récurrence pour $x_1, \cdots, x_{n-1}, x_n \in G$

$$x_1 * \cdots * x_n = (x_1 * \cdots * x_{n-1}) * x_n.$$

**Lemme 1. 1.** *Soient $x_1, \cdots, x_{n-1}, x_n \in G$ $(n \geqslant 3)$. Tout produit formé avec $x_1, \cdots, x_{n-1}, x_n$ dans cet ordre et "des parenthèses placées de façon quelconque" (par exemple $(x_1 * (x_2 * x_3)) * ((x_4 * x_5) * (x_6 * x_7))$ si $n = 7$) est égal à $x_1 * \cdots * x_n$.*

**Démonstration.**

Pour $n = 3$, c'est la définition de l'associativité. Soit $p$ tel que le résultat soit vrai pour $n \leqslant p$ (hypothèse de récurrence), et démontrons qu'il est vrai pour $n = p+1$. Soient donc $x_1, \cdots, x_{p+1} \in G$ et $x$ un produit formé avec $x_1, \cdots, x_{p+1}$. Il existe $q$ $(1 \leqslant q \leqslant p)$ tel que $x = y * z$ où $y$ est un produit formé avec $x_1, \cdots, x_q$ et $z$ est un produit formé avec $x_{q+1}, \cdots, x_{p+1}$ (par exemple

$$(x_1 * (x_2 * x_3)) * ((x_4 * x_5) * (x_6 * x_7)) = y * z$$

avec $y = x_1 * (x_2 * x_3)$ et $z = (x_4 * x_5) * (x_6 * x_7)$). Deux cas sont à distinguer :

**1er cas :** $q = p$. Dans ce cas, $z = x_{p+1}$. Alors $y$ est formé avec $x_1, \cdots, x_p$ donc par l'hypothèse de récurrence $y = x_1 * \cdots * x_p$. Donc

$$x = y * z = (x_1 * \cdots * x_p) * x_{p+1} = x_1 * \cdots * x_p * x_{p+1}.$$

**2ème cas :** $q \leqslant p - 1$. Alors par l'hypothèse de récurrence $y = x_1 * \cdots * x_q$ et $z = x_{q+1} * \cdots * x_{p+1}$ $(q + 1 \leqslant p)$. Donc

$$\begin{aligned} x &= (x_1 * \cdots * x_q) * (x_{q+1} * \cdots * x_{p+1}) \\ &= (x_1 * \cdots * x_q) * ((x_{q+1} * \cdots * x_p) * x_{p+1}) \\ &= ((x_1 * \cdots * x_q) * (x_{q+1} * \cdots * x_p)) * x_{p+1} \\ &= (x_1 * \cdots * x_q * x_{q+1} * \cdots * x_p) * x_{p+1} \\ &= x_1 * \cdots * x_q * x_{q+1} * \cdots * x_p * x_{p+1}. \end{aligned}$$

Ainsi le résultat est vrai pour $n = p + 1$. Il est donc vrai pour tout $n \in \mathbb{N}$, $n \geqslant 3$ par récurrence. □

### 1.3.2 Règles.

Si $G$ est un groupe noté multiplicativement (*resp.* additivement), on pose pour $x \in G$

$$x^n = \begin{cases} \underbrace{x \cdots x}_{n \text{ fois}} & \text{si } n \geqslant 1 \\ e & \text{si } n = 0 \end{cases} \tag{1.1}$$

(*resp.*

$$nx = \begin{cases} \underbrace{x + \cdots + x}_{n \text{ fois}} & \text{si } n \geqslant 1 \\ 0 & \text{si } n = 0 \end{cases}). \tag{1.2}$$

Pour $n \leqslant 0$, on note

$$x^{-n} = (x^{-1})^n \quad ( \textit{resp. } (-n)x = n(-x)). \tag{1.3}$$

Alors

$$x^{-n}x^n = \underbrace{x^{-1} \cdots x^{-1}}_{n \text{ fois}} \underbrace{x \cdots x}_{n \text{ fois}} = e$$

donc

$$x^{-n} = (x^n)^{-1}$$

et de même dans le cas d'un groupe additif

$$(-n)x = -(nx).$$

**Lemme 1. 2.** *Soit $G$ un groupe multiplicatif. Pour tous $x \in G$ , $p, q \in \mathbb{Z}$ on a*

$$x^p x^q = x^{p+q} \quad (1.4)$$
$$(x^p)^q = x^{pq} \quad (1.5)$$

**Démonstration.**

Si $p, q \in \mathbb{N}$ les égalités sont immédiates. Si $p, q < 0$,

$$x^p x^q = x^{-|p|} x^{-|q|} = (x^{-1})^{|p|}(x^{-1})^{|q|} = (x^{-1})^{|p|+|q|} = x^{-(|p|+|q|)} = x^{p+q}.$$

$$(x^p)^q = (x^{-|p|})^{-|q|} = (x^{|p|})^{|q|} = x^{|pq|} = x^{pq}.$$

Sinon, on peut supposer $p < 0$ et $q > 0$, le raisonnement étant analogue si $p > 0$ et $q < 0$. Alors

$$x^p x^q = \underbrace{x^{-1} \cdots x^{-1}}_{|p| \text{ fois}} \underbrace{x \cdots x}_{q \text{ fois}} = x^{-|p|+q} = x^{p+q}$$

$$(x^p)^q = (x^{-|p|})^q = ((x^{-1})^{|p|})^q = (x^{-1})^{|p|q} = x^{pq}.$$

□

Le lemme précédent donne dans le cas d'un groupe additif :

$$px + qx = (p+q)x \quad , \quad q(px) = (qp)x \quad \forall x \in G, \ p, q \in \mathbb{Z}.$$

On notera que dans un groupe multiplicatif

$$(xy)^{-1} = y^{-1}x^{-1} \quad \forall x, y \in G.$$

De plus, $(xy)^2 = xyxy$ n'est en général pas égal à $x^2y^2$. On a en effet :

$$\begin{aligned}(xy)^2 = x^2y^2 \quad &\Leftrightarrow \quad xyxy = xxyy \\ &\Leftrightarrow \quad yx = xy.\end{aligned}$$

## 1.4 Anneaux, corps.

### 1.4.1 Définition.

**Définition 1. 2.** *On appelle anneau un ensemble $\mathcal{A}$ muni de deux lois de compositions internes $+$ et $\cdot$ ayant les propriétés suivantes.*

*(i) Muni de la loi $+$, $\mathcal{A}$ est un groupe commutatif.*

*(ii) La loi $\cdot$ est associative*

$$x \cdot (y \cdot z) = (x \cdot y) \cdot z \quad \forall\, x, y, z \in \mathcal{A}$$

*et distributive par rapport à l'addition*

$$x \cdot (y + z) = x \cdot y + x \cdot z \quad \forall\, x, y, z \in \mathcal{A}$$

$$(y + z) \cdot x = y \cdot x + z \cdot x \quad \forall\, x, y, z \in \mathcal{A}.$$

*On note 0 l'élément neutre de l'addition. S'il existe dans $\mathcal{A}$ un élément neutre pour la multiplication $\cdot$ , on le note 1 ou $1_{\mathcal{A}}$ et on dit que l'anneau est unitaire. Si la multiplication est commutative, on dit que l'anneau est commutatif. On appelle corps un anneau unitaire tel que $1 \neq 0$ et dans lequel tout élément non nul possède un inverse pour la multiplication, i.e.*

$$\forall x \in \mathcal{A} \setminus \{0\}, \quad \exists y \in \mathcal{A} \quad x \cdot y = y \cdot x = 1.$$

*L'inverse $y$ de $x \in \mathcal{A} \setminus \{0\}$ se note alors $x^{-1}$.*
*On notera le plus souvent $xy$ au lieu de $x \cdot y$.*

**Remarques.** (i) Si $\mathcal{A}$ est un anneau, on a $0 \cdot x = x \cdot 0 = 0 \ \forall x \in \mathcal{A}$. En effet, $0 \cdot x = (0+0) \cdot x = 0 \cdot x + 0 \cdot x$ donc $0 \cdot x = 0 \cdot x - 0 \cdot x = 0$. De même pour $x \cdot 0$. On a aussi $(-x) \cdot y = -(x \cdot y)$ pour tous $x, y \in \mathcal{A}$ puisque $x \cdot y + (-x) \cdot y = (x - x) \cdot y = 0 \cdot y = 0$. De même, $x \cdot (-y) = -(x \cdot y)$.
(ii) Un anneau *nul* $\mathcal{A} = \{0\}$ est unitaire puisque $0 \cdot 0 = 0$, et l'on a dans ce cas $1 = 0$. Réciproquement, si $\mathcal{A}$ est un anneau unitaire tel que $1 = 0$, alors $\mathcal{A} = \{0\}$. En effet, si $1 = 0$, on a $x = 1 \cdot x = 0 \cdot x = 0$ pour tout $x \in \mathcal{A}$, donc $\mathcal{A} = \{0\}$. Si $\mathcal{A}$ est un anneau unitaire non nul, on a ainsi $1 \neq 0$. Un anneau unitaire non nul possède donc au moins deux éléments, à savoir 0 et 1.
(iii) Un anneau unitaire $\mathcal{A}$ est un corps si et seulement si l'ensemble $\mathcal{A}^* = \mathcal{A} \setminus \{0\}$ est un groupe multiplicatif.
(iv) Si $\mathcal{A}$ est un anneau unitaire et $n \in \mathbb{Z}$, $x \in \mathcal{A}$, l'élément $nx \in \mathcal{A}$ est défini par (1.2) et (1.3). On a alors

$$nx = (n\,1)x = x(n\,1)\,.$$

En effet, pour $n = 0$, $nx = 0$ et $n\,1 = 0$ ; pour $n > 0$ , c'est une application de la distributivité; enfin pour $n < 0$,

$$nx = (-|n|)x = -(|n|x) = \begin{cases} -((|n|\,1)x) = (-(|n|\,1))x = ((-|n|)\,1)x = (n\,1)x \\ -(x(|n|\,1)) = x(-(|n|\,1)) = x((-|n|)\,1) = x(n\,1)\,. \end{cases}$$

### 1.4.2 Exemples.

(i) $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ sont des corps commutatifs. $\mathbb{Z}$ est un anneau commutatif unitaire.

(ii) L'ensemble $M_n(\mathbb{C})$ (*resp.* $M_n(\mathbb{R}), M_n(\mathbb{Q}), M_n(\mathbb{Z})$) des matrices $n \times n$ à coefficients dans $\mathbb{C}$ (*resp.* réels, rationnels, entiers), muni de l'addition et de la multiplication des matrices, est un anneau unitaire. Il est *non commutatif* pour $n \geqslant 2$. L'élément neutre de la multiplication est la matrice $I_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$. On la note simplement $I$ si $n$ est précisé par le contexte, et on l'appelle *matrice identité*.

(iii) L'ensemble $\mathbb{H}$ des matrices de la forme $\begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \in M_2(\mathbb{C})$ avec $\alpha, \beta \in \mathbb{C}$ est un corps non commutatif appelé *corps des quaternions* (voir ex. 1.20).

(iv) $n\mathbb{Z}$ ($n \geqslant 2$) est un anneau commutatif *non unitaire*.

(v) L'ensemble $\mathcal{A} = \mathcal{C}(\mathbb{R})$ des fonctions continues $f : \mathbb{R} \to \mathbb{R}$, muni de l'addition et de la multiplication habituelles définies pour tous $f, g \in \mathcal{A}$ et tout $x \in \mathbb{R}$ par

$$\begin{aligned} (f+g)(x) &= f(x) + g(x) \\ (fg)(x) &= f(x)g(x) \end{aligned}$$

est un anneau commutatif unitaire l'élément unité étant la fonction constante 1.

(vi) L'ensemble $\mathcal{A} = \mathcal{C}_0(\mathbb{R})$ des fonctions continues $f : \mathbb{R} \to \mathbb{R}$ telles que $f(0) = 0$, muni de l'addition et de la multiplication habituelles, est un anneau commutatif *non unitaire.* S'il existait une fonction $\phi \in \mathcal{A}$ telle que $\phi f = f\ \forall f \in \mathcal{A}$, on aurait en effet, en prenant en particulier pour $f$ la fonction $f(x) = x$, $\phi(x)x = x\ \forall x \in \mathbb{R}$, donc $\phi(x) = 1$ pour $x \neq 0$. On aurait ainsi $\phi(x) = 0$ pour $x = 0$ et $\phi(x) = 1$ pour $x \neq 0$. Cela contredirait la continuité de la fonction $\phi$. Donc une telle fonction n'existe pas.

### 1.4.3 Identités remarquables dans un anneau.

**Formule du binôme.**

Soit $\mathcal{A}$ un anneau et $x, y$ deux éléments de $\mathcal{A}$ qui *commutent*, *i.e.* $xy = yx$. Alors

$$(x+y)^n = \sum_{k=0}^{n} C_n^k\, x^{n-k} y^k \quad \forall n \geqslant 1\,, \tag{1.6}$$

où

$$C_n^k = \frac{n(n-1)\cdots(n-k+1)}{k!} \tag{1.7}$$

pour $1 \leqslant k \leqslant n$ et $C_n^0 = 1$. Pour $n = 1$, la formule se réduit à $x + y = x + y$. Supposons la formule vraie pour $n = p$. Alors pour $n = p + 1$,

$$\begin{aligned}
(x+y)^{p+1} &= (x+y)(x+y)^p \\
&= (x+y)\sum_{k=0}^{p} C_p^k\, x^{p-k} y^k \\
&= \sum_{k=0}^{p} C_p^k\, x^{p-k+1} y^k + \sum_{k=0}^{p} C_p^k\, x^{p-k} y^{k+1} \\
&= \sum_{k=0}^{p} C_p^k\, x^{p-k+1} y^k + \sum_{h=1}^{p+1} C_p^{h-1}\, x^{p-h+1} y^h \\
&= x^{p+1} + \sum_{k=1}^{p} (C_p^k + C_p^{k-1})\, x^{p-k+1} y^k + y^{p+1} \\
&= x^{p+1} + \sum_{k=1}^{p} C_{p+1}^k\, x^{p-k+1} y^k + y^{p+1} \\
&= \sum_{k=0}^{p+1} C_{p+1}^k\, x^{p-k+1} y^k
\end{aligned}$$

donc la formule est vraie pour $n = p + 1$. Par récurrence, elle est donc vraie pour tout $n \geqslant 1$.

**Identité $x^n - y^n$.**

Soit $\mathcal{A}$ un anneau et $x, y$ deux éléments de $\mathcal{A}$ qui *commutent.* Alors

$$x^n - y^n = (x-y)(x^{n-1} + x^{n-2}y + \cdots + xy^{n-2} + y^{n-1}) \quad \forall n \geqslant 2. \tag{1.8}$$

La vérification de cette formule est immédiate par distributivité.

## 1.5 Groupe symétrique $\mathcal{S}_n$.

### 1.5.1 Définitions.

**Théorème 1. 1.** *Soit $X$ un ensemble et* Bij$(X)$ *l'ensemble des bijections de $X$ sur lui-même.* Bij$(X)$ *muni de la loi $\circ$ de composition des applications est un groupe.*

**Démonstration.**

On sait que la composée de deux bijections est une bijection. La loi $\circ$ de composition des applications est une donc une loi interne sur Bij$(X)$. De plus si $h : X \to Y$, $g : Y \to Z$, $f : Z \to T$ sont 3 applications entre les ensembles $X, Y, Z, T$ on a $f \circ (g \circ h) = (f \circ g) \circ h$ puisque pour tout $x \in X$ :

$$(f \circ (g \circ h))(x) = f((g \circ h)(x)) = f(g(h(x))) = (f \circ g)(h(x)) = ((f \circ g) \circ h)(x).$$

La loi $\circ$ est donc associative. L'application identique *Id* sur $X$ est clairement élément neutre. Pour $f \in$ Bij$(X)$, la bijection réciproque de $f$ est élément symétrique de $f$ pour la loi $\circ$. □

**Définition 1. 3.** *Pour $n \in \mathbb{N}^*$, on appelle groupe symétrique sur $n$ éléments et on note $\mathcal{S}_n$ le groupe des bijections de l'ensemble $\{1, \dots, n\}$ muni de la loi $\circ$. Les éléments de $\mathcal{S}_n$ s'appellent les permutations de $\{1, \dots, n\}$.*

Pour $s, t \in \mathcal{S}_n$ on note simplement $st$ au lieu de $s \circ t$. Un élément $s \in \mathcal{S}_n$ se note $\begin{pmatrix} 1 & \dots & n \\ s(1) & \dots & s(n) \end{pmatrix}$. On appelle *permutation circulaire sur $p$ éléments* ou *$p$-cycle* ($1 \leqslant p \leqslant n$) un élément $s \in \mathcal{S}_n$ pour lequel il existe une suite de $p$ entiers deux-à-deux distincts $a_1, a_2, \dots, a_p$ appartenant à $\{1, \dots, n\}$ tels que $s(a_1) = a_2, \dots, s(a_{p-1}) = a_p$, $s(a_p) = a_1$ et $s(k) = k$ pour $k \neq a_1, \dots, a_p$. On note alors $s = (a_1, \dots, a_p)$. Un 2-cycle s'appelle une *transposition*.

### 1.5.2 Non-commutativité.

Pour $n \geqslant 3$, le groupe $\mathcal{S}_n$ est *non commutatif.* On a en effet : $(1,2)(2,3) = (1,2,3)$ alors que $(2,3)(1,2) = (1,3,2)$.

### 1.5.3 Une formule fondamentale.

**Théorème 1. 2.** *Soit $(a_1, \dots, a_p) \in \mathcal{S}_n$ un $p$-cycle. Pour toute permutation $s \in \mathcal{S}_n$ on a :*

$$s(a_1, \dots, a_p)s^{-1} = (s(a_1), \dots, s(a_p)). \tag{1.9}$$

**Démonstration.**

Pour tout $k \in \{1, \dots, n\} \setminus \{s(a_1), \dots, s(a_p)\}$, on a $s^{-1}(k) \notin \{a_1, \dots, a_p\}$, donc $(a_1, \dots, a_p)(s^{-1}(k)) = s^{-1}(k)$, et $(s(a_1, \dots, a_p)s^{-1})\ (k) = s(s^{-1}(k)) = k$. D'autre part, on a pour $1 \leqslant j \leqslant p$

$$\left(s(a_1, \dots, a_p)s^{-1}\right)\ (s(a_j)) = (s(a_1, \dots, a_p))\ (a_j) = \begin{cases} s(a_{j+1}) & \text{si } 1 \leqslant j < p \\ s(a_1) & \text{si } j = p \end{cases}.$$

□

## 1.6 Groupe et anneau $\mathbb{Z} / n\mathbb{Z}$.

### 1.6.1 Ensemble quotient.

Une *relation d'équivalence* sur un ensemble $X$ est une relation $\mathcal{R}$ sur $X$ vérifiant les 3 propriétés suivantes :

$$\begin{array}{rl} \text{(Réflexivité)} & x\mathcal{R}x \quad \forall x \in X \\ \text{(Symétrie)} & x\mathcal{R}y \Leftrightarrow y\mathcal{R}z \quad \forall x, y \in X \\ \text{(Transitivité)} & x\mathcal{R}y \text{ et } y\mathcal{R}z \Rightarrow x\mathcal{R}z \quad \forall x, y, z \in X \end{array}$$

Pour $x \in X$ , $[x] = \{y \in X;\ x\mathcal{R}y\}$ est la *classe d'équivalence de $x$ modulo $\mathcal{R}$*. Les classes d'équivalence modulo $\mathcal{R}$ forment une partition de $X$, et l'ensemble des classes d'équivalence (modulo $\mathcal{R}$) est appelé *ensemble quotient (modulo $\mathcal{R}$)*.

### 1.6.2 Équivalence modulo $n$ dans $\mathbb{Z}$.

On prend ici $X = \mathbb{Z}$ et on définit pour $n \in \mathbb{N}^*$ la relation $\mathcal{R}$ par $x\mathcal{R}y \Leftrightarrow x - y \in n\mathbb{Z}$. On voit facilement que $\mathcal{R}$ est une relation d'équivalence sur $\mathbb{Z}$. On note au lieu de $x\mathcal{R}y$ : $x \equiv y \pmod n$. L'ensemble quotient est noté $\mathbb{Z} / n\mathbb{Z}$; ses éléments sont les $n$ classes distinctes $[r]\ 0 \leqslant r \leqslant n-1$.

**Théorème 1. 3.** *(i) $\mathbb{Z} / n\mathbb{Z}$ est un anneau commutatif unitaire.*
*(ii) L'anneau $\mathbb{Z} / n\mathbb{Z}$ est un corps si et seulement si $n$ est premier.*

**Démonstration.**
(i) Il faut d'abord définir la somme et le produit de deux classes. Soient $[x], [y] \in \mathbb{Z} / n\mathbb{Z}$. Les éléments $x, y \in \mathbb{Z}$ sont des *représentants* des classes respectives $[x], [y]$. Si $x', y'$ sont des autres représentants des classes respectives $[x], [y]$, on aura $x' \equiv x \pmod n$, et $y' \equiv y \pmod n$, donc $x' = x + kn$ et $y' = y + \ell n$ avec $k, \ell \in \mathbb{Z}$. Alors $x' + y' = x + y + (k + \ell)n$ donc $x' + y' \equiv x + y \pmod n$ et $[x' + y'] = [x + y]$. Ainsi la classe $[x + y]$ *ne dépend pas des représentants $x, y$ utilisés pour* $[x], [y]$. On la note $[x] + [y]$. Cela définit une application $+ : \mathbb{Z} / n\mathbb{Z} \times \mathbb{Z} / n\mathbb{Z} \to \mathbb{Z} / n\mathbb{Z}$. De même, $x'y' = (x + kn)(y + \ell n) = xy + (ky + \ell x + k\ell n)n$ donc $x'y' \equiv xy \pmod n$, $[x'y'] = [xy]$ et la classe $[xy]$ ne dépend pas des représentants $x, y$ utilisés pour $[x], [y]$. On la note $[x] \cdot [y]$ et cela définit une application $\cdot : \mathbb{Z} / n\mathbb{Z} \times \mathbb{Z} / n\mathbb{Z} \to \mathbb{Z} / n\mathbb{Z}$. Il reste maintenant à vérifier les différents axiomes. L'addition est associative puisque pour tous $x, y, z \in \mathbb{Z}$

$$[x] + ([y] + [z]) = [x] + [y + z] = [x + y + z] = [x + y] + [z] = ([x] + [y]) + [z].$$

On vérifie de même que : $[0]$ est élément neutre pour l'addition, $[-x]$ est l'opposé de $[x]$, et $[x] + [y] = [y] + [x]\ \forall x, y \in \mathbb{Z}$. Donc $\mathbb{Z} / n\mathbb{Z}$ est un groupe commutatif pour l'addition. Les propriétés relatives à la multiplication se vérifient de la même façon.
(ii) L'anneau $\mathbb{Z} / n\mathbb{Z}$ est un corps si et seulement si $[0] \neq [1]$, *i.e.* $n \geqslant 2$, et tout élément non nul est inversible pour la multiplication, *i.e.*

$$\forall [x] \neq [0],\ \exists\, [y] \qquad [x][y] = [1].$$

Comme $[x][y] = [xy]$, cette dernière condition équivaut pour $n \geqslant 2$ à

$$\forall x \in \{1, 2, \dots, n-1\},\ \exists\, y, k \in \mathbb{Z} \qquad xy - kn = 1, \qquad .$$

ou encore d'après l'identité de Bezout

$$\forall x \in \{1, 2, \ldots, n-1\}, \quad n \curlywedge x = 1$$

(en notant $n \curlywedge x$ le PGCD de $n$ et $x$) ce qui signifie que $n$ est premier. □

Nous noterons $\mathbb{Z}_n$ le *groupe additif* $\mathbb{Z} / n\mathbb{Z}$.

# 1.7 Sous-groupes.

## 1.7.1 Définition.

**Définition 1. 4.** *Soit $G$ un groupe (noté multiplicativement). Un sous-ensemble $H$ de $G$ est appelée un sous-groupe de $G$ s'il possède les 3 propriétés suivantes :*
- *L'élément neutre $e$ de $G$ appartient à $H$.*
- *$H$ est stable pour la multiplication, i.e. $xy \in H \ \forall x, y \in H$.*
- *$H$ est stable pour le passage à l'inverse, i.e. $x^{-1} \in H \ \forall x \in H$.*

Si $H$ est un sous-groupe du groupe $G$, il est immédiat que la multiplication dans $G$ induit sur $H$ une structure de groupe.

## 1.7.2 Exemples.

(i) $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ sont des sous-groupes du groupe additif $\mathbb{C}$.

(ii) $\mathbb{R}^*_+$ est un sous-groupe du groupe multiplicatif $\mathbb{R}^*$.

(iii) $\mathbb{T}$ et $\mu_n$ sont des sous-groupes du groupe multiplicatif $\mathbb{C}^*$.

(iv) $K$ désignant un corps commutatif, l'ensemble $SL(n, K)$ des matrices $n \times n$ à coefficients dans $K$ et de déterminant 1 est un sous-groupe du groupe $GL(n, K)$ des matrices $n \times n$ inversibles à coefficients dans $K$ muni de la multiplication matricielle.

## 1.7.3 Sous-groupes additifs de $\mathbb{Z}$.

**Théorème 1. 4.** *Les sous-groupes additifs de $\mathbb{Z}$ sont les sous-ensembles de la forme $n\mathbb{Z}$ avec $n \in \mathbb{N}$.*

**Démonstration.**

Soit $n \in \mathbb{N}$ et $H = n\mathbb{Z} = \{\ldots, -3n, -2n, -n, 0, n, 2n, 3n, \ldots\}$. $H$ est un sous-groupe additif de $\mathbb{Z}$ puisque :
- L'élément neutre $0 = n \cdot 0$ de $\mathbb{Z}$ appartient à $H$.
- $H$ est stable pour l'addition;

$$x + y = np + nq = n(p + q) \in H \quad \forall x = np, y = nq \in H, p, q \in \mathbb{Z};$$

- $H$ est stable pour le passage à l'opposé,

$$-x = -(np) = n(-p) \in H \quad \forall x = np \in H, p \in \mathbb{Z}.$$

Réciproquement, soit $H$ un sous-groupe additif quelconque de $\mathbb{Z}$, et montrons qu'il existe un $n \in \mathbb{N}$ unique tel que $H = n\mathbb{Z}$. Si $H = \{0\}$, $n = 0$ est le seul entier de $\mathbb{N}$ qui convient. Supposons donc $H \neq \{0\}$. Notons que si $n > 0$, $n$ est le plus petit élément strictement positif de $n\mathbb{Z}$. S'il existe $n \in \mathbb{N}$ tel que $H = n\mathbb{Z}$, on a nécessairement $n > 0$ et $n$ est le plus petit élément strictement positif de $H$.

D'où l'unicité de $n$. Pour l'existence, considérons l'ensemble $H_+ = \{h \in H; h > 0\}$. Comme $H \neq \{0\}$, il existe $h \in H$ non nul et changeant éventuellement $h$ en $-h$, on peut supposer $h > 0$. Donc $H_+$ est non vide. $H_+$ est un sous-ensemble non vide de $\mathbb{N}$ donc il possède un *plus petit élément* $n$. On a $n > 0$ puisque $n \in H_+$. Nous allons montrer que $H = n\mathbb{Z}$. D'abord $H \supset n\mathbb{Z}$. En effet, $n \in H$ et $H$ est un sous-groupe. Il faut maintenant montrer que $H \subset n\mathbb{Z}$. Soit $x \in H$ et montrons que $x \in n\mathbb{Z}$. On peut supposer $x \geqslant 0$. Il existe $p \in \mathbb{Z}$ et $r \in \mathbb{N}, 0 \leqslant r < n$, tels que $x = np + r$. Comme $np \in n\mathbb{Z} \subset H$ on a $r = x - np \in H$ puisque $H$ est un sous-groupe additif. Si l'on avait $r > 0$, on aurait $r \in H_+$. Or ce n'est pas possible car $r < n$ contredirait la définition de $n$. Donc $r = 0$ et $x = np \in n\mathbb{Z}$. □

### 1.7.4 Sous-groupes additifs de $\mathbb{R}$.

**Théorème 1. 5.** *Un sous-groupe additif de $\mathbb{R}$ est soit dense soit de la forme $a\mathbb{Z}$ avec $a \in \mathbb{R}, a \geqslant 0$.*

**Démonstration.**

Soit $H$ un sous-groupe additif de $\mathbb{R}$. Si $H = \{0\}$, $H = a\mathbb{Z}$ avec $a = 0$. Supposons donc $H \neq \{0\}$. Alors il existe $h \in H$ non nul. Changeant éventuellement $h$ en $-h$, on peut supposer $h > 0$. Donc l'ensemble $H_+ = \{h \in H; h > 0\}$ est non vide. $H_+$ est un sous-ensemble non vide de $\mathbb{R}$ minoré par 0, donc il possède une borne inférieure $a \geqslant 0$. Deux cas peuvent se produire.

**1er cas :** $a = 0$. Montrons que dans ce cas $H$ est dense dans $\mathbb{R}$, *i.e.* pour tout $x \in \mathbb{R}$ et tout $\varepsilon > 0$, il existe $h \in H$ tel que $|h - x| < \varepsilon$. Changeant éventuellement $x$ en $-x$ et $h$ en $-h$, on peut supposer $x \geqslant 0$. Soient donc $x \geqslant 0$ et $\varepsilon > 0$ quelconques. Comme $a = 0$, il existe $k \in H_+$ tel que $0 < k < \varepsilon$. Si $N$ est la partie entière de $\frac{x}{k}$, on a $N \leqslant \frac{x}{k} < N + 1$, ce qui s'écrit $Nk \leqslant x < (N+1)k$, donc $0 \leqslant x - Nk < k < \varepsilon$. L'élément $h = Nk \in H$ répond à la question.

**2ème cas :** $a > 0$. Montrons que dans ce cas $H = a\mathbb{Z}$ (en particulier H est non dense dans $\mathbb{R}$).

i) D'abord $a$ appartient à $H$. Supposons en effet $a \notin H$. Comme $a < 2a$, il existe, par définition d'une borne inférieure, un $h \in H_+$ tel que $a \leqslant h < 2a$, et l'on a $a < h$ puisque $a \notin H$. De même, il existe $k \in H_+$ tel que $a < k < h$. Alors $h - k \in H_+$ et $h - k < a$, ce qui est impossible puisque $a$ est la borne inférieure de $H_+$.

ii) On a $a\mathbb{Z} \subset H$ puisque $a \in H$. Il faut maintenant montrer que $H \subset a\mathbb{Z}$. Soit $x \in H$ et montrons que $x \in a\mathbb{Z}$. On peut supposer $x \geqslant 0$. On a $x = ma + r$ où $m$ est la partie entière de $\frac{x}{a}$ et $0 \leqslant r < a$. Alors $r = x - ma \in H$. Si l'on avait $r > 0$, on aurait $r \in H_+$. Or ce n'est pas possible car $r < a$ contredirait la définition de $a$. Donc $r = 0$ et $x = ma \in a\mathbb{Z}$. □

## 1.8 Sous-groupe engendré par une partie.

**Lemme 1. 3.** *Soit $G$ un groupe. L'intersection $H = \bigcap_{i \in I} H_i$ d'une famille non vide quelconque $(H_i)_{i \in I}$ de sous-groupes de $G$ est un sous-groupe de $G$.*

**Démonstration.**

Par hypothèse, $I \neq \emptyset$. $e \in H$ car $e \in H_i \; \forall i \in I$; pour tous $x, y \in H$, on a $x, y \in H_i \; \forall i \in I$ donc $xy \in H_i \; \forall i \in I$ et $xy \in H$. Enfin, pour tout $x \in H$, on a $x \in H_i \; \forall i \in I$ donc $x^{-1} \in H_i \; \forall i \in I$ et $x^{-1} \in H$. □

**Exemple.** Soit $G$ un groupe et $X$ une partie de $G$. L'intersection de la famille de tous les sous-groupes de $G$ contenant $X$ est un sous-groupe. Il est immédiat que c'est le plus petit sous-groupe (pour la relation d'inclusion) de $G$ contenant $X$. On l'appelle sous-groupe engendré par $X$ et on le note $\langle X\rangle$. On a donc par définition :

$$\langle X\rangle = \bigcap_{\substack{H\supset X\\ H\text{sous-groupe}}} H$$

Lorsque $X$ est réduit à un élément $x\in G$, $X=\{x\}$, on note simplement $\langle x\rangle$ au lieu de $H=\langle\{x\}\rangle$ et on dit simplement que $\langle x\rangle$ est le sous-groupe engendré par $x$.

**Lemme 1. 4.** *Soit $G$ un groupe et $x\in G$. On a :*

$$\langle x\rangle = \{x^n;\, n\in\mathbb{Z}\}.$$

**Démonstration.**

Le sous-groupe $\langle x\rangle$ de $G$ contient $x$ donc contient aussi $x^n$ $\forall n\in\mathbb{Z}$, donc $\langle x\rangle\supset\{x^n;\, n\in\mathbb{Z}\}$. Or on voit facilement que $\{x^n;\, n\in\mathbb{Z}\}$ est un sous-groupe de $G$, et comme il contient $x$, il contient le sous-groupe engendré par $x$. D'où l'égalité. □

**Définition 1. 5.** *Un groupe $G$ est dit monogène s'il existe $x\in G$ tel que $G=\langle x\rangle$. Un groupe monogène fini est dit cyclique.*

**Lemme 1. 5.** *Soit $G$ un groupe, $X$ une partie non vide de $G$ et $\langle X\rangle$ le sous-groupe engendré par $X$. On a :*

$$\langle X\rangle = \left\{x\in G;\ \exists p\in\mathbb{N}^*, x=x_1\dots x_p,\quad x_i\in X \text{ ou } x_i^{-1}\in X\quad \forall i,\ 1\leqslant i\leqslant p\right\}.$$

**Démonstration.**

Soit

$$A = \left\{x\in G;\ \exists p\in\mathbb{N}^*, x=x_1\dots x_p,\quad x_i\in X \text{ ou } x_i^{-1}\in X\quad \forall i,\ 1\leqslant i\leqslant p\right\}.$$

On a $X\subset A\subset G$. Il est immédiat que tout sous-groupe de $G$ contenant $X$ contient aussi $A$. Donc $A\subset\langle X\rangle$. Pour montrer qu'on a l'égalité, il suffit de montrer que $A$ est un sous-groupe de $G$. D'abord comme $X\neq\varnothing$, il existe $a\in X$ et donc $e=aa^{-1}\in A$. Ensuite si $x=x_1\dots x_p$ ($x_i\in X$ ou $x_i^{-1}\in X$ $\forall i,\ 1\leqslant i\leqslant p$) et $y=y_1\dots y_q$ ($y_j\in X$ ou $y_j^{-1}\in X$ $\forall j,\ 1\leqslant j\leqslant q$) sont deux éléments de $A$, on a $xy=x_1\dots x_py_1\dots y_q=z_1\dots z_{p+q}$ avec $z_k=x_k$ pour $1\leqslant k\leqslant p$ et $z_k=y_{k-p}$ pour $p+1\leqslant k\leqslant p+q$. Comme $z_k\in X$ ou $z_k^{-1}\in X$ $\forall k,\ 1\leqslant k\leqslant p+q$, on a bien $xy\in A$. Enfin $x^{-1}=(x_1\dots x_p)^{-1}=x_p^{-1}\dots x_1^{-1}=s_1\dots s_p$ avec $s_i=x_{p+1-i}^{-1}$, donc $s_i\in X$ ou $s_i^{-1}\in X$ $\forall i,\ 1\leqslant i\leqslant p$ et $x^{-1}\in A$. Cela prouve que $A$ est un sous-groupe de $G$.

□

## 1.9 Homomorphismes de groupes.

### 1.9.1 Définition.

**Définition 1. 6.** *Soient $G, G'$ deux groupes. Une application $f : G\longrightarrow G'$ est appelée un homorphisme de groupes si elle possède la propriété suivante :*

$$f(xy)=f(x)f(y)\quad\forall\, x,y\in G.$$

*Si en plus $f$ est bijectif, on dit que $f$ est un isomorphisme de $G$ sur $G'$, et on note alors $G\cong G'$.*

### 1.9.2 Exemples.

Les applications suivantes sont des homomorphismes de groupes.
(i) $f : \mathbb{R} \longrightarrow \mathbb{R}_+^* \quad x \mapsto f(x) = e^x$.
(ii) $f : \mathbb{R} \longrightarrow \mathbb{T} \quad x \mapsto f(x) = e^{ix} = \cos x + i \sin x$.
(iii) $f : \mathbb{C} \longrightarrow \mathbb{C}^* \quad z \mapsto f(z) = e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$.
(iv) $f : GL(n, \mathbb{R}) \longrightarrow \mathbb{R}^* \quad A \mapsto f(A) = \det A$ où $\det A$ désigne le déterminant de $A$.

### 1.9.3 Propriétés immédiates.

**Proposition 1. 1.** *Soient $G, G'$ deux groupes et $f : G \longrightarrow G'$ un homorphisme de groupes. Alors*
*(i)* $f(e) = e'$ *($e, e'$ désignent les éléments neutres de $G, G'$ respectivement).*
*(ii)* $f(x^{-1}) = (f(x))^{-1} \quad \forall x \in G$.

**Démonstration.**
(i) On a $f(e) = f(ee) = f(e)f(e)$, donc $e' = f(e)$ par multiplication par $(f(e))^{-1}$.
(ii) Pour tout $x \in G$, on a $f(x^{-1})f(x) = f(x^{-1}x) = f(e) = e'$ donc $f(x^{-1}) = (f(x))^{-1}$. □

**Proposition 1. 2.** *(i) Soient $G, G', G''$ des groupes et $f : G \longrightarrow G'$, $g : G' \longrightarrow G''$ des homorphismes de groupes. Alors l'application composée $h = g \circ f : G \longrightarrow G''$ est un homorphisme de groupes.*
*(ii) Si $f : G \longrightarrow G'$ est un isomorphisme de groupes, la bijection réciproque $g = f^{-1} : G' \longrightarrow G$ est aussi un isomorphisme de groupes.*

**Démonstration.**
(i) On a pour tous $x, y \in G$

$$(g \circ f)(xy) = g(f(xy)) = g(f(x)f(y)) = g(f(x))g(f(y)) = (g \circ f)(x)\,(g \circ f)(y).$$

(ii) Pour tous $x', y' \in G'$ on a en posant $x = f^{-1}(x'), y = f^{-1}(y')$ : $f(xy) = f(x)f(y) = x'y'$ donc $f^{-1}(x'y') = xy = f^{-1}(x)f^{-1}(y')$. Ainsi $f^{-1}$ est un homomorphisme de groupes, donc un isomorphisme puisque bijectif. □

### 1.9.4 Automorphismes, sous-groupes distingués.

**Définition 1. 7.** *Soit $G$ un groupe. On appelle automorphisme de $G$ un isomorphisme de $G$ sur lui-même. On note* Aut$(G)$ *l'ensemble des automorphismes de $G$.*

**Proposition 1. 3.** Aut$(G)$ *est un groupe pour la loi* $\circ$.

**Démonstration.**
On va montrer que Aut$(G)$ est un sous-groupe du groupe Bij$(G)$ des bijections de $G$ sur lui-même.
• Par définition Aut$(G) \subset$ Bij$(G)$.
• L'application identique $Id \in$ Bij$(G)$ est un homomorphisme, donc appartient à Aut$(G)$.

• Pour tous $f, g \in \mathrm{Aut}\,(G)$, la bijection $f \circ g$ est un homomorphisme puisque $f$ et $g$ sont des homomorphismes, donc $f \circ g \in \mathrm{Aut}\,(G)$.
• Pour tout $f \in \mathrm{Aut}\,(G)$, la bijection réciproque $f^{-1}$ est un isomorphisme de $G$ sur $G$, donc $f^{-1} \in \mathrm{Aut}\,(G)$. □

**Automorphismes intérieurs.**

**Proposition 1. 4.** *Soit $G$ un groupe. Pour tout $x \in G$, l'application* $\mathrm{Int}\,(x)$ *de $G$ dans $G$ définie par* $\mathrm{Int}\,(x)(y) = xyx^{-1} \quad \forall y \in G$ *est un automorphisme de $G$. On appelle* $\mathrm{Int}\,(x)$ *l'automorphisme intérieur de $G$ défini par $x$.*

**Démonstration.**
Pour tout $x \in G$, $\mathrm{Int}\,(x)$ est bijectif puisque

$$\forall z, y \in G, \quad z = \mathrm{Int}\,(x)(y) \Leftrightarrow z = xyx^{-1} \Leftrightarrow y = x^{-1}zx,$$

*i.e.* tout $z \in G$ possède un antécédent unique. D'autre part $\mathrm{Int}\,(x)$ est un homomorphisme puisque

$$\mathrm{Int}\,(x)(yz) = xyzx^{-1} = xyx^{-1}xzx^{-1} = \mathrm{Int}\,(x)(y)\,\mathrm{Int}\,(x)(z)\ \forall y, z \in G.$$

□

**Remarque.** Si $G$ est commutatif, le seul automorphisme intérieur est l'identité.

**Sous-groupe distingué.**

**Définition 1. 8.** *Un sous-groupe $H$ d'un groupe $G$ est dit distingué dans $G$ (notation : $H \lhd G$) s'il est stable par tous les automorphismes intérieurs de $G$, i.e.*

$$\mathrm{Int}\,(x)(H) \subset H \quad \forall x \in G$$

*ce qui s'écrit aussi*

$$xHx^{-1} \subset H \quad \forall x \in G \tag{1.10}$$

**Remarque.** Si $G$ est commutatif, tout sous-groupe de $G$ est distingué dans $G$.

**Proposition 1. 5.** *Un sous-groupe $H$ d'un groupe $G$ est distingué dans $G$ si et seulement si*

$$xHx^{-1} = H \quad \forall x \in G \tag{1.11}$$

**Démonstration.**
La condition (1.11) implique la condition (1.10). Montrons que la condition (1.10) implique la condition (1.11). Soit donc $H$ un sous-groupe de $G$ vérifiant (1.10). Pour montrer que $H$ vérifie la condition (1.11), il suffit de montrer que

$$H \subset xHx^{-1} \quad \forall x \in G \tag{1.12}$$

Soit donc $a \in G$ quelconque. La condition (1.10) appliquée avec $x = a^{-1}$ donne

$$a^{-1}Ha \subset H$$

ce qui s'écrit encore en multipliant à gauche par $a$ et à droite par $a^{-1}$ :

$$H \subset aHa^{-1}.$$

Comme $a$ est arbitraire, (1.12) est démontrée. □

**Classes de conjugaison.**

**Définition 1. 9.** *Soit $G$ un groupe. On appelle conjugaison dans $G$ la relation définie dans $G$ par*

$$x\mathcal{R}y \Leftrightarrow \exists a \in G \quad y = axa^{-1}.$$

On vérifie facilement que la conjugaison est une relation d'équivalence dans $G$. Une classe d'équivalence est appelée une *classe de conjugaison de $G$*.

**Centre d'un groupe.**

**Définition 1. 10.** *On appelle centre d'un groupe $G$ le sous-ensemble*

$$Z(G) = \{a \in G\,;\, ax = xa \quad \forall x \in G\}.$$

**Proposition 1. 6.** *Le centre $Z(G)$ d'un groupe $G$ est un sous-groupe distingué de $G$.*

**Démonstration.**

D'abord $Z(G)$ est un sous-groupe de $G$ :
- Par définition $Z(G) \subset G$.
- L'élément neutre $e$ de $G$ appartient au centre $Z(G)$.
- Pour tous $a, b \in Z(G)$, on a $abx = axb = xab \quad \forall\, x \in G$, donc $ab \in Z(G)$.
- Pour tout $a \in Z(G)$ et tout $x \in G$, on a $ax = xa$ d'où, successivement $x = a^{-1}xa$ et $xa^{-1} = a^{-1}x$. Donc $a^{-1} \in Z(G)$.

Ensuite, on a pour tout $a \in Z(G)$ et tout $x \in G$

$$xax^{-1} = axx^{-1} = a \in Z(G)$$

donc $Z(G)$ est distingué dans $G$. □

### 1.9.5 Propriétés des homomorphismes, image, noyau.

Rappelons que si $f : X \longrightarrow Y$ est une application d'un ensemble $X$ dans un ensemble $Y$, on appelle *image réciproque* d'une partie $B$ de $Y$, et on note $f^{-1}(B)$, le sous-ensemble de $X$

$$f^{-1}(B) = \{x \in X;\ f(x) \in B\}.$$

On appelle *image* d'une partie $A$ de $X$, et on note $f(A)$, le sous-ensemble de $Y$

$$f(A) = \{y \in Y;\ \exists x \in A \quad y = f(x)\}.$$

**Théorème 1. 6.** *Soit $f : G \longrightarrow G'$ un homomorphisme de groupes et $e, e'$ les éléments neutres de $G, G'$ respectivement.*
*(i) $Ker\, f = f^{-1}(\{e'\})$ est un sous-groupe distingué de $G$.*
*(ii) $f$ est injectif si et seulement si $Ker\, f = \{e\}$.*
*(iii) $Im\ f = f(G)$ est un sous-groupe de $G'$.*
*(iv) Pour tout sous-groupe $H$ de $G$, $f(H)$ est un sous-groupe de $G'$. Si $H$ est distingué dans $G$, $f(H)$ est distingué dans $f(G)$.*
*(v) Pour tout sous-groupe $H'$ de $G'$, $f^{-1}(H')$ est un sous-groupe de $G$. Si $H'$ est distingué dans $G'$, $f^{-1}(H')$ est distingué dans $G$.*

**Démonstration.**

(i) D'abord $\operatorname{Ker} f$ est un sous-groupe de $G$:

• Par définition $\operatorname{Ker} f \subset G$.

• D'après la Prop. 1.1 $f(e) = e'$ donc $e \in \operatorname{Ker} f$.

• Pour tous $x, y \in \operatorname{Ker} f$, on a $f(xy) = f(x)f(y) = e'e' = e'$, donc $xy \in \operatorname{Ker} f$.

• Pour tout $x \in \operatorname{Ker} f$, $f(x^{-1}) = (f(x))^{-1} = (e')^{-1} = e'$ donc $x^{-1} \in \operatorname{Ker} f$.

Ensuite, pour vérifier que $\operatorname{Ker} f$ est distingué dans $G$ il faut vérifier que $yxy^{-1} \in \operatorname{Ker} f\ \forall x \in \operatorname{Ker} f, \forall y \in G$. Or

$$f(yxy^{-1}) = f(y)f(x)f(y^{-1}) = f(y)f(y^{-1}) = e'$$

puisque $f(x) = e$ et $f(y^{-1}) = (f(y))^{-1}$. Donc $f(yxy^{-1}) = e'$ et $yxy^{-1} \in \operatorname{Ker} f$.

(ii) Supposons $f$ injectif. Soit $x \in \operatorname{Ker} f$. On a $f(x) = f(e)$ donc $x = e$. Cela prouve que $\operatorname{Ker} f = \{e\}$. Réciproquement, supposons $\operatorname{Ker} f = \{e\}$. Soit $x, y \in G$ tels que $f(x) = f(y)$. Alors $f(xy^{-1}) = f(x)(f(y))^{-1} = e$, donc $xy^{-1} \in \operatorname{Ker} f$. Comme $\operatorname{Ker} f = \{e\}, xy^{-1} = e$ et donc $x = y$. Cela prouve que $f$ est injective. D'où l'équivalence : $f$ injective $\Leftrightarrow \operatorname{Ker} f = \{e\}$.

(iii)

• Par définition $f(G) \subset G'$.

• $e' = f(e) \in f(G)$.

• Pour tous $x', y' \in f(G)$, il existe $x, y \in G$ tels que $x' = f(x)$ et $y' = f(y)$. On a alors $x'y' = f(x)f(y) = f(xy)$, donc $x'y' \in f(G)$.

• Pour tout $x' \in f(G)$, il existe $x \in H$ tel que $x' = f(x)$ et alors $(x')^{-1} = (f(x))^{-1} = f(x^{-1}) \in f(G)$.

(iv) D'abord $f(H)$ est un sous-groupe de $G'$ : il suffit de reprendre en le généralisant le raisonnement de (iii).

• Par définition $f(H) \subset G'$.

• $e' = f(e) \in f(H)$.

• Pour tous $x', y' \in f(H)$, il existe $x, y \in H$ tels que $x' = f(x)$ et $y' = f(y)$. On a alors $x'y' = f(x)f(y) = f(xy)$. Or $xy \in H$ puisque $H$ est un sous-groupe de $G$, donc $x'y' \in f(H)$.

• Pour tout $x' \in f(H)$, il existe $x \in H$ tel que $x' = f(x)$ et alors $(x')^{-1} = (f(x))^{-1} = f(x^{-1}) \in f(H)$ puisque $x^{-1} \in H$.

Ensuite, pour vérifier que $f(H)$ est distingué dans $f(G)$ il faut vérifier que $y'x'(y')^{-1} \in f(H)\ \forall x' \in f(H), \forall y' \in f(G)$. Il existe $x \in H$ et $y \in G$ tels que $x' = f(x)$ et $y' = f(y)$. Alors

$$y'x'(y')^{-1} = f(y)f(x)(f(y))^{-1} = f(y)f(x)f(y^{-1}) = f(yxy^{-1}).$$

Or $yxy^{-1} \in H$ puisque $x \in H$ et que $H$ est distingué dans $G$. Donc $y'x'(y')^{-1} \in f(H)$ d'où le résultat.

(v) il suffit de reprendre en le généralisant le raisonnement de (i). D'abord $f^{-1}(H')$ est un sous-groupe de $G$:

• Par définition $f^{-1}(H') \subset G$.

• On a $f(e) = e'$. Or $e' \in H'$ puisque $H'$ est un sous-groupe de $G'$. Donc $e \in f^{-1}(H')$.

• Pour tous $x, y \in f^{-1}(H')$, on a $f(x) \in H'$ et $f(y) \in H'$. Alors $f(xy) = f(x)f(y) \in H'$ puisque $H'$ est un sous-groupe de $G'$. Donc $xy \in f^{-1}(H')$.

• Pour tout $x \in f^{-1}(H')$, $f(x^{-1}) = (f(x))^{-1} \in H'$ puisque $f(x) \in H'$ et que $H'$ est un sous-groupe de $G'$. Donc $x^{-1} \in f^{-1}(H')$.

Ensuite, pour vérifier que $f^{-1}(H')$ est distingué dans $G$ il faut vérifier que $yxy^{-1} \in f^{-1}(H')\ \forall x \in f^{-1}(H'), \forall y \in G$. Or

$$f(yxy^{-1}) = f(y)f(x)f(y^{-1}) = f(y)f(x)(f(y))^{-1} \in H'$$

puisque $f(x) \in H'$, $f(y) \in G'$ et $H'$ est distingué dans $G'$. D'où le résultat. □

**Définition 1. 11.** *Soit $f : G \longrightarrow G'$ un homomorphisme de groupes. Le sous-groupe distingué $Ker\, f = f^{-1}(\{e'\})$ de $G$ s'appelle le noyau de $f$, et le sous-groupe $Im\ f = f(G)$ de $G'$ s'appelle l'image de $f$.*

## 1.10 Classes à gauche, à droite, groupe quotient.

### 1.10.1 Équivalence à gauche modulo $H$.

**Définition 1. 12.** *Soit $G$ un groupe et $H$ un sous-groupe de $G$. On appelle équivalence à gauche modulo $H$ la relation définie dans $G$ par*

$$x \equiv y \pmod H \Leftrightarrow x^{-1}y \in H.$$

On vérifie facilement que l'équivalence à gauche modulo $H$ est une relation d'équivalence dans $G$. La classe d'équivalence d'un élément $x \in G$ est le sous-ensemble $[x] = xH = \{y \in G;\ \exists h \in H \quad y = xh\}$; on l'appelle *classe de $x$ à gauche modulo $H$*. L'ensemble quotient se note $G/H$. On l'appelle *espace homogène* (à gauche) modulo $H$.

**Proposition 1. 7.** *Pour tous $x, y \in G$, on a :*

$$x \equiv y \pmod H \Rightarrow ax \equiv ay \pmod H \quad \forall a \in G.$$

*On dit que l'équivalence à gauche modulo $H$ est compatible à gauche avec la multiplication de $G$.*

**Démonstration.**
$(ax)^{-1}ay = x^{-1}a^{-1}ay = x^{-1}y \in H.$ □

### 1.10.2 Équivalence à droite modulo $H$.

On définit de même l'équivalence à droite modulo $H$ par

$$x \equiv_d y \pmod H \Leftrightarrow yx^{-1} \in H.$$

L'équivalence à droite modulo $H$ est aussi une relation d'équivalence dans $G$. La classe d'équivalence d'un élément $x \in G$ est le sous-ensemble $[x]_d = Hx = \{y \in G;\ \exists h \in H \quad y = hx\}$; on l'appelle *classe de $x$ à droite modulo $H$*. L'ensemble quotient est appelé *espace homogène à droite* modulo $H$. Pour tous $x, y \in G$, on a :

$$x \equiv_d y \pmod H \Rightarrow xa \equiv_d ya \pmod H \quad \forall a \in G.$$

On dit que l'équivalence à droite modulo $H$ est compatible à droite avec la multiplication de $G$.

### 1.10.3 Cas d'un sous-groupe distingué : groupe quotient.

**Proposition 1. 8.** *Soit $G$ un groupe et $H$ un sous-groupe de $G$. $H$ est distingué dans $G$ si et seulement si pour tout $x \in G$ la classe à gauche $xH$ modulo $H$ coïncide avec la classe à droite $Hx$ modulo $H$. Cela signifie aussi que les deux relations d'équivalence à gauche et à droite modulo $H$ coïncident.*

**Démonstration.**

$$\begin{aligned} H \lhd G &\Leftrightarrow xHx^{-1} \subset H \quad \forall x \in G \\ &\Leftrightarrow xHx^{-1} = H \quad \forall x \in G \text{ (Prop. 1.5)} \\ &\Leftrightarrow xH = Hx \quad \forall x \in G \end{aligned}$$

□

**Théorème 1. 7.** *Soit $G$ un groupe et $H$ un sous-groupe distingué de $G$. Alors $G/H$ est un groupe appelé groupe quotient de $G$ par $H$. L'application $\pi : G \longrightarrow G/H$ définie par $\pi(x) = [x] = xH$ est un homomorphisme de groupes surjectif appelé projection canonique.*

**Démonstration.**

Il faut d'abord définir le produit de deux classes. Soient $[x], [y] \in G/H$. Les éléments $x, y \in G$ sont des *représentants* des classes respectives $[x], [y]$. Si $x', y'$ sont des autres représentants des classes respectives $[x], [y]$, on aura $x' \equiv x \pmod H$, et $y' \equiv y \pmod H$, donc $x' = xh$ et $y' = yk$ avec $h, k \in H$. Alors $x'y' = xhyk = xyy^{-1}hyk$. Or $y^{-1}hy \in H$ puisque $H$ est distingué. Donc $y^{-1}hyk \in H$ et alors $x'y' \equiv xy \pmod H$, et $[x'y'] = [xy]$. Ainsi la classe $[xy]$ *ne dépend pas des représentants $x, y$ utilisés pour* $[x], [y]$. On la note $[x][y]$. Cela définit une multiplication $\cdot : G/H \times G/H \to G/H$. La multiplication est associative puisque pour tous $x, y, z \in G$

$$[x]([y][z]) = [x][yz] = [xyz] = [xy][z] = ([x][y])[z].$$

On vérifie de même que $[e]$ est élément neutre, et que $[x^{-1}]$ est l'inverse de $[x]$. Enfin, l'application $\pi : G \longrightarrow G/H$ définie par $\pi(x) = [x] = xH$ est surjective par définition de $G/H$. C'est un homomorphisme par définition de la multiplication dans $G/H$. □

### 1.10.4 Exemple.

Pour $G = \mathbb{Z}$ et $H = n\mathbb{Z}$, $G/H$ est le groupe additif $\mathbb{Z}_n = \mathbb{Z} / n\mathbb{Z}$ déjà défini.

### 1.10.5 Théorème de Lagrange.

**Théorème 1. 8 (Théorème de Lagrange).** *Soit $G$ un groupe et $H$ un sous-groupe de $G$. Si $G$ est un groupe fini, alors l'ordre de $H$ divise l'ordre de $G$.*

**Démonstration.**

Si $G$ est d'ordre fini $n$, $H$ est aussi d'ordre fini. Notons $p$ l'ordre de $H$. Les classes à gauche modulo $H$ forment une partition de $G$; elles sont en nombre fini $q$. Elles ont toutes le même cardinal $p$ que $H$, puisque pour tout $x \in G$ l'application $h \mapsto xh$ de $H$ sur la classe $xH$ est une bijection. On a donc $n = pq$. □

## 1.11 Décomposition canonique d'un homomorphisme de groupes.

**Théorème 1. 9.** *Soit* $f : G \longrightarrow G'$ *un homomorphisme de groupes et* $\pi : G \longrightarrow G/Ker f$ *la projection canonique. Il existe un homomorphisme unique*

$$\tilde{f} : G/Ker f \longrightarrow G'$$

*rendant le diagramme suivant commutatif:*

$$\begin{array}{ccc} G & \xrightarrow{f} & G' \\ \pi\downarrow & \nearrow & \tilde{f} \\ G/Ker f & & \end{array}$$

*L'homomorphisme* $\tilde{f}$ *est injectif et c'est un isomorphisme de* $G/Ker f$ *sur* $Im\ f$.

**Démonstration.**

Pour $x \in G$, $\pi(x) = [x] = x \operatorname{Ker} f$ est la classe à gauche de $x$ modulo $\operatorname{Ker} f$ (elle coïncide avec la classe à droite puisque $\operatorname{Ker} f$ est distingué). S'il existe une application $\tilde{f} : G/\operatorname{Ker} f \longrightarrow G'$ rendant le diagramme commutatif, on aura nécessairement pour tout $x \in G$ :

$$\tilde{f}([x]) = \tilde{f}(\pi(x)) = f(x).$$

Cette expression montre que si une telle application existe, elle est unique. Montrons qu'elle existe. Soit $\pi(x) = [x] = x \operatorname{Ker} f \in G/\operatorname{Ker} f$. L'élément $x \in G$ est un représentant de la classe $[x]$. Si $y \in G$ est un autre représentant de cette classe, on a $y \equiv x \pmod{\operatorname{Ker} f}$, *i.e.* il existe $h \in \operatorname{Ker} f$ tel que $y = xh$. On a alors $f(y) = f(xh) = f(x)f(h) = f(x)$. L'élément $f(x) \in G'$ ne dépend donc pas du représentant utilisé pour $[x]$. On définit alors une application $\tilde{f} : G/\operatorname{Ker} f \longrightarrow G'$ en posant $\tilde{f}([x]) = f(x)$. Par définition, cette application rend le diagramme commutatif.

L'application $\tilde{f}$ est un homomorphisme de groupes. On a en effet pour tous $x, y \in G$ :

$$\tilde{f}([x][y]) = \tilde{f}([xy]) = f(xy) = f(x)f(y) = \tilde{f}([x])\tilde{f}([y]).$$

L'homomorphisme $\tilde{f}$ est injectif. On a en effet $\operatorname{Ker} \tilde{f} = \{[e]\}$ puisque

$$\tilde{f}([x]) = e' \Leftrightarrow f(x) = e' \Leftrightarrow x \in \operatorname{Ker} f \Leftrightarrow [x] = [e].$$

Enfin, l'image de $\tilde{f}$ est la même que celle de $f$, donc $\tilde{f}$ est une bijection de $G/H$ sur $\operatorname{Im} f$. □

## 1.12 Structure des groupes monogènes.

### 1.12.1 Théorème de structure des groupes monogènes.

**Théorème 1. 10.** *Soit* $G$ *un groupe monogène. Alors :*

$$G \cong \begin{cases} \mathbb{Z} & \textit{si } G \textit{ est infini} \\ \mathbb{Z}_n & \textit{si } G \textit{ est fini d'ordre } n. \end{cases}$$

**Démonstration.**

Soit $x \in G$ tel que $G = \langle x \rangle$. L'application $f : \mathbb{Z} \to G$ définie par $f(x) = x^n$ est un homomorphisme de groupes, et $f$ est surjectif. Le noyau $H = Ker\, f$ est un sous-groupe de $\mathbb{Z}$. Si $H = \{0\}$, $f$ est injectif, donc est un isomorphisme de $\mathbb{Z}$ sur $G$. Sinon, $H$ est de la forme $n\mathbb{Z}$ avec $n \in \mathbb{N}^*$ et la décomposition canonique de $f$ donne un isomorphisme $\tilde{f}$ de $\mathbb{Z}_n$ sur $G$. □

### 1.12.2 Ordre d'un élément.

**Définition 1. 13.** *On appelle ordre d'un élément $x$ d'un groupe $G$ l'ordre du sous-groupe $\langle x \rangle$ de $G$ engendré par $x$ si ce sous-groupe est fini. Sinon, on dit que $x$ est d'ordre infini.*

**Théorème 1. 11.** *Soit $G$ un groupe et $x \in G$.*
*(i) $x$ est d'ordre fini si et seulement si $H_+ = \{k \in \mathbb{N}^*; x^k = e\} \neq \varnothing$. L'ordre $d$ de $x$ est alors le plus petit élément de $H_+$, et l'on a :*

$$\{k \in \mathbb{Z}; x^k = e\} = d\mathbb{Z}, \qquad \langle x \rangle = \{e, x, x^2, \cdots, x^{d-1}\}.$$

*(ii) Si $G$ est un groupe d'ordre fini $n$, tout $x \in G$ est d'ordre fini et l'ordre $d$ de $x$ divise $n$.*
*(iii) Si $G$ est un groupe d'ordre fini $n$, alors $x^n = e\, \forall x \in G$.*

**Démonstration.**

(i) Soit $f : \mathbb{Z} \to \langle x \rangle$ défini par $f(k) = x^k\ \forall k \in \mathbb{Z}$. $f$ est un homomorphisme, et il est surjectif puisque $\langle x \rangle = \{x^n;\ n \in \mathbb{Z}\}$. Or

$$\begin{aligned} H_+ = \varnothing &\Leftrightarrow\ x^k \neq e \quad \forall k \in \mathbb{N}^* \\ &\Leftrightarrow\ x^k \neq e \quad \forall k \in \mathbb{Z}^* \\ &\Leftrightarrow\ \operatorname{Ker} f = \{0\}. \end{aligned}$$

Donc d'après le Th.1.10, $x$ est d'ordre fini si et seulement si $H_+ \neq \varnothing$. Dans ce cas, si $d$ désigne l'ordre de $x$, on a $\operatorname{Ker} f = d\mathbb{Z}$ et la décomposition canonique de $f$ donne un isomorphisme $\tilde{f}$ de $\mathbb{Z}_d$ sur $\langle x \rangle$. Comme $\tilde{f}([r]) = x^r$ pour $0 \leqslant r < d$ et que $\tilde{f}$ est surjectif,

$$\langle x \rangle = \{e, x, x^2, \cdots, x^{d-1}\}.$$

(ii) $H = \langle x \rangle$ est un sous-groupe de $G$, donc l'ordre $d$ de $H$ est fini, et divise $n$ d'après le Théorème de Lagrange.
(iii) Soit $x \in G$ quelconque. D'après (ii), l'ordre $d$ de $x$ divise $n$ : $n = kd$. Alors $x^n = x^{kd} = (x^d)^k = e$. □

**Remarque.** Dans le cas où le groupe $G$ est noté additivement (avec élément neutre 0) et non pas multiplicativement, les parties (i) et (iii) du Théorème 1.11 se lisent :
(i)' $x$ est d'ordre fini si et seulement si $H_+ = \{k \in \mathbb{N}^*; kx = 0\} \neq \varnothing$. L'ordre $d$ de $x$ est alors le plus petit élément de $H_+$, et l'on a :

$$\{k \in \mathbb{Z}; kx = 0\} = d\mathbb{Z}, \qquad \langle x \rangle = \{e, x, 2x, \cdots, (d-1)x\}.$$

(iii)' Si $G$ est un groupe d'ordre fini $n$, alors $nx = 0\ \forall x \in G$.

**Exemple.** Dans le groupe additif $G = \mathbb{Z}_6$, les divers éléments ont les ordres suivants :

| élément | ordre |
|---|---|
| [0] | 1 |
| [1] | 6 |
| [2] | 3 |
| [3] | 2 |
| [4] | 3 |
| [5] | 6 |

### 1.12.3 Sous-groupes des groupes monogènes.

**Théorème 1. 12.** *Tout sous-groupe d'un groupe monogène est monogène.*

**Démonstration.**
Soit $G = \langle x \rangle$ et $H$ un sous-groupe de $G$. L'application $f : \mathbb{Z} \to G$ définie par $f(k) = x^k$ est un homomorphisme de groupes, et $f$ est surjectif puisque $G$ est engendré par $x$. Comme $H$ est un sous-groupe de $G$, l'image réciproque $f^{-1}(H)$ de $H$ par $f$ est un sous-groupe de $\mathbb{Z}$, donc il existe $p \in \mathbb{N}$ tel que $f^{-1}(H) = p\mathbb{Z}$. Soit $y \in H$ arbitraire. Comme $x$ engendre $G$, il existe $n \in \mathbb{Z}$ tel que $y = x^n$. On a alors $n \in f^{-1}(H)$ puisque $y \in H$, donc $n = pq$ avec $q \in \mathbb{Z}$. Cela implique $y = x^{pq} = (x^p)^q \in \langle x^p \rangle$. Comme $y \in H$ est arbitraire, $H \subset \langle x^p \rangle$. Mais $p \in p\mathbb{Z} = f^{-1}(H)$ donc $x^p \in H$ ce qui implique $\langle x^p \rangle \subset H$, d'où l'égalité. □

### 1.12.4 Sous-groupes des groupes cycliques.

**Théorème 1. 13.** *Soit $G = \langle x \rangle$ un groupe cyclique d'ordre $n$. Pour tout diviseur $d$ de $n$, il existe un unique sous-groupe d'ordre $d$ de $G$. Ce sous-groupe est $H_d = \langle x^p \rangle$ avec $p = \frac{n}{d}$. On a aussi $H_d = \{y \in G;\ y^d = e\}$.*

**Démonstration.**
Soit $p = \frac{n}{d}$ . L'ordre de l'élément $x^p$ de $G$ est $d$ et le sous-groupe de $G$ engendré par $x^p$ est

$$H_d = \langle x^p \rangle = \{e, x^p, \dots, x^{(d-1)p}\}.$$

C'est un groupe cyclique d'ordre $d$.

Montrons que $H_d$ est le seul sous-groupe d'ordre $d$ de $G$. Soit $H$ un sous-groupe d'ordre $d$ de $G$. On sait (Th.1.12) que $H$ est monogène, donc il existe $r \in \{0, \dots, n-1\}$ tel que $H = \langle x^r \rangle$. Comme $H$ est d'ordre $d$, il en est de même de $x^r$, donc $x^{rd} = e$. Comme $x$ est d'ordre $n$, cela implique que $n$ divise $rd$: $rd = kn$ avec $k \in \mathbb{N}$. On a alors :

$$x^r = x^{\frac{rd}{d}} = x^{\frac{nk}{d}} = x^{pk} \in \langle x^p \rangle = H_d$$

donc $H = \langle x^r \rangle \subset H_d$. Mais $H$ et $H_d$ ont le même ordre $d$, donc on a l'égalité $H = H_d$.

Enfin, $K = \{y \in G;\ y^d = e\}$ est un sous-groupe de $G$, et $H_d \subset K$ puisque $H_d$ est d'ordre $d$. Or $K$ est monogène, donc il existe $s \in \{0, \dots, n-1\}$ tel que $K = \langle x^s \rangle$. On a $x^s \in K$, donc $x^{sd} = e$, ce qui implique comme ci-dessus $x^s \in H_d$. Alors $K \subset H_d$, d'où l'égalité . □

**Remarque.** Le Théorème 1.13 s'énonce dans le cas du groupe cyclique *additif* $\mathbb{Z}_n$ ($n \in \mathbb{N}^*$) : pour tout diviseur $d$ de $n$, il existe un unique sous-groupe d'ordre $d$ de $\mathbb{Z}_n$. Ce sous-groupe est $H_d = \langle [p] \rangle$ avec $p = \frac{n}{d}$, et l'on a aussi $H_d = \{[k];\ k \in \mathbb{Z};\ d[k] = [0]\}$.

# 1.13 Produit direct.

**Théorème 1. 14.** *Soient $G_1$ et $G_2$ deux groupes. L'ensemble produit cartésien $G_1 \times G_2 = \{(x,y); x \in G_1, y \in G_2\}$ muni de la loi interne*

$$(x,y) \cdot (x',y') = (xx', yy') \tag{1.13}$$

*est un groupe. Le groupe $G_1 \times G_2$ est commutatif si et seulement si $G_1$ et $G_2$ le sont.*

**Démonstration.**

Il est clair que la loi est bien une loi interne sur l'ensemble produit cartésien $G_1 \times G_2$. Elle est associative car pour tous $(x,y), (x',y'), (x'',y'') \in G_1 \times G_2$, on a

$$\begin{aligned}
(x,y) \cdot ((x',y') \cdot (x'',y'')) &= (x,y) \cdot (x'x'', y'y'') \\
&= (xx'x'', yy'y'') \\
&= (xx', yy') \cdot (x'', y'') \\
&= ((x,y) \cdot (x',y')) \cdot (x'',y'').
\end{aligned}$$

Si $e_1$ et $e_2$ désignent les éléments neutres de $G_1$ et $G_2$ respectivement, il est immédiat que le couple $(e_1, e_2)$ est élément neutre de $G_1 \times G_2$. Pour tout $(x,y) \in G_1 \times G_2$, l'élément $(x^{-1}, y^{-1}) \in G_1 \times G_2$ vérifie

$$(x,y) \cdot (x^{-1}, y^{-1}) = (x^{-1}, y^{-1}) \cdot (x,y) = (e_1, e_2).$$

$G_1 \times G_2$ est donc un groupe. Enfin on a :

$$\begin{aligned}
G_1 \times G_2 \text{ commutatif} &\Leftrightarrow (x,y) \cdot (x',y') = (x',y') \cdot (x,y) \\
&\qquad \forall x, x' \in G_1\ \forall y, y' \in G_2 \\
&\Leftrightarrow (xx', yy') = (x'x, y'y)\ \forall x, x' \in G_1\ \forall y, y' \in G_2 \\
&\Leftrightarrow xx' = x'x \text{ et } yy' = y'y\ \forall x, x' \in G_1\ \forall y, y' \in G_2 \\
&\Leftrightarrow G_1 \text{ commutatif et } G_2 \text{ commutatif}.
\end{aligned}$$

□

**Définition 1. 14.** *Le groupe $G_1 \times G_2$ avec la loi (1.13) est appelé le produit direct des deux groupes $G_1$ et $G_2$.*

**Théorème 1. 15.** *Soient $p, q$ deux entiers $> 0$ premiers entre eux : $p \curlywedge q = 1$. Alors $\mathbb{Z}_{pq}$ est isomorphe au produit direct $\mathbb{Z}_p \times \mathbb{Z}_q$.*

**Démonstration.**

Considérons l'application $f : \mathbb{Z} \to \mathbb{Z}_p \times \mathbb{Z}_q$ définie par $f(x) = ([x]_p, [x]_q)$, où $[x]_p$ et $[x]_q$ désignent respectivement les classes de $x$ modulo $p$ et modulo $q$. On a

$$\begin{aligned}
f(x + x') &= ([x + x']_p, [x + x']_q) \\
&= ([x]_p + [x']_p, [x]_q + [x']_q) \\
&= ([x]_p, [x]_q) + ([x']_p, [x']_q) \\
&= f(x) + f(x')
\end{aligned}$$

donc $f$ est un homomorphisme de groupes. Le noyau $\operatorname{Ker} f$ est formé des $x \in \mathbb{Z}$ tels que $([x]_p, [x]_q) = ([0],[0])$, *i.e.* $x$ est divisible par $p$ et $x$ est divisible par $q$. Mais $p$ et $q$ sont premier entre eux, donc cela équivaut à dire que $x$ est divisible par $pq$, *i.e.* $x \in pq\mathbb{Z}$. Ainsi $\operatorname{Ker} f = pq\mathbb{Z}$. La décomposition canonique de l'homomorphisme $f$ donne alors un homomorphisme injectif

$$\tilde{f} : \mathbb{Z}/pq\mathbb{Z} = \mathbb{Z}_{pq} \longrightarrow \mathbb{Z}_p \times \mathbb{Z}_q. \tag{1.14}$$

Comme $\mathbb{Z}_{pq}$ et $\mathbb{Z}_p \times \mathbb{Z}_q$ ont le même cardinal $pq$, $\tilde{f}$ est bijectif, donc c'est un isomorphisme. □

**Corollaire (Lemme chinois).** *Soient $p, q$ deux entiers $> 0$ premiers entre eux. Pour tous $k, \ell \in \mathbb{Z}$, il existe $n \in \mathbb{Z}$ tel que $n \equiv k \pmod p$ et $n \equiv \ell \pmod q$.*

**Démonstration.**

Considérons l'élément $([k]_p, [\ell]_q)$ de $\mathbb{Z}_p \times \mathbb{Z}_q$. Comme l'application $\tilde{f}$ de (1.14) est bijective, il existe $n \in \mathbb{Z}$ tel que $([k]_p, [\ell]_q) = \tilde{f}([n])$. Or $\tilde{f}([n]) = ([n]_p, [n]_q)$, donc $[k]_p = [n]_p$ et $[\ell]_q = [n]_q$, *i.e.* $n \equiv k \pmod p$ et $n \equiv \ell \pmod q$. □

# 1.14 Notions analogues dans les anneaux.

## 1.14.1 Sous-anneaux.

**Définition 1. 15.** *Soit $\mathcal{A}$ un anneau. On appelle sous-anneau de $\mathcal{A}$ tout sous-ensemble $\mathcal{B} \subset \mathcal{A}$ ayant les propriétés suivantes.*

*(i) $\mathcal{B}$ est un sous-groupe additif de $\mathcal{A}$.*

*(ii) $\mathcal{B}$ est stable pour la multiplication dans $\mathcal{A}$, i.e. $xy \in \mathcal{B}$ pour tous $x, y \in \mathcal{B}$.*

Si $\mathcal{B}$ est un sous-anneau de l'anneau $\mathcal{A}$, il est immédiat que l'addition et la multiplication dans $\mathcal{A}$ induisent sur $\mathcal{B}$ une structure d'anneau.

## 1.14.2 Sous-corps.

**Définition 1. 16.** *Soit $K$ un corps. On appelle sous-corps de $K$ tout sous-ensemble $L \subset K$ ayant les propriétés suivantes.*

*(i) $L$ est un sous-anneau de $K$.*

*(ii) $1 \in L$.*

*(iii) $x^{-1} \in L \quad \forall x \in L,\ x \neq 0$.*

Si $L$ est un sous-corps du corps $K$, il est immédiat que l'addition et la multiplication dans $K$ induisent sur $L$ une structure de corps.

## 1.14.3 Idéaux.

**Définition 1. 17.** *Soit $\mathcal{A}$ un anneau quelconque. On appelle idéal bilatère de $\mathcal{A}$ tout sous-ensemble $\mathcal{I} \subset \mathcal{A}$ ayant les propriétés suivantes.*

*(i) $\mathcal{I}$ est un sous-groupe additif de $\mathcal{A}$.*

*(ii) $ax \in \mathcal{I}$ et $xa \in \mathcal{I}$ pour tout $a \in \mathcal{A}$ et tout $x \in \mathcal{I}$.*

*Lorsque l'anneau $\mathcal{A}$ est commutatif, on dit simplement "idéal" au lieu de "idéal bilatère".*

**Exemple.** Les idéaux de $\mathbb{Z}$ sont les sous-groupes additifs $n\mathbb{Z}$, $n \in \mathbb{N}$.

**Proposition 1. 9.** *Les seuls idéaux bilatères d'un corps $K$ sont $\{0\}$ et $K$.*

**Démonstration.**
Soit $\mathcal{I}$ un idéal bilatère du corps $K$. Si $\mathcal{I} \neq \{0\}$, il existe $x \in \mathcal{I}$, $x \neq 0$. Alors $1 = x^{-1}x \in \mathcal{I}$. On a alors $y = y \cdot 1 \in \mathcal{I}$ pour tout $y \in K$, donc $\mathcal{I} = K$. □

**Proposition 1. 10.** *Soit $\mathcal{A}$ un anneau commutatif unitaire. Pour tout $a \in \mathcal{A}$, le sous-ensemble $a\mathcal{A} = \{ax;\ x \in \mathcal{A}\}$ de $\mathcal{A}$ est un idéal de $\mathcal{A}$. C'est le plus petit idéal (pour l'inclusion) contenant $a$.*

**Démonstration.**
On a $0 = a \cdot 0 \in a\mathcal{A}$, $ax + ay = a(x+y) \in a\mathcal{A}, -ax = a(-x) \in a\mathcal{A}$, $\forall x, y \in \mathcal{A}$, donc $a\mathcal{A}$ est un sous-groupe additif de $\mathcal{A}$. De plus pour tous $x, y \in \mathcal{A}$, $(ax)y = a(xy) \in a\mathcal{A}$ donc $a\mathcal{A}$ est un idéal de l'anneau commutatif $\mathcal{A}$. Il contient l'élément $a$ puisque $a = a \cdot 1 \in a\mathcal{A}$. Enfin, si $\mathcal{I}$ est un idéal quelconque de $\mathcal{A}$ contenant l'élément $a$, on aura $ax \in \mathcal{I}$ pour tout $x \in \mathcal{A}$, donc $a\mathcal{A} \subset \mathcal{I}$. □

**Définition 1. 18.** *Soit $\mathcal{A}$ un anneau commutatif unitaire et $a \in \mathcal{A}$. L'idéal $a\mathcal{A}$ est appelé l'idéal engendré par $a$ et est noté $(a)$. Un idéal $\mathcal{I}$ de $\mathcal{A}$ est dit principal s'il existe $a \in \mathcal{A}$ tel que $\mathcal{I} = (a)$. L'anneau $\mathcal{A}$ est dit principal si tout idéal est principal.*

**Exemples.**
(i) $\mathbb{Z}$ est principal.
(ii) Soit $K$ un corps commutatif. L'anneau $K[X]$ des polynômes à l'indéterminée $X$ est principal. D'une façon précise, on a :

**Théorème 1. 16.** *Soit $\mathcal{I}$ un idéal de $K[X]$.*
*(i) Il existe $P \in K[X]$ tel que $\mathcal{I} = (P)$.*
*(ii) $P$ est déterminé de façon unique par $\mathcal{I}$, à un facteur multiplicatif constant $\neq 0$ près.*

**Démonstration.**
(i) • Si $\mathcal{I} = \{0\}$, $P = 0$, et (ii) est dans ce cas trivial.
• Supposons $\mathcal{I} \neq \{0\}$. Il existe dans $\mathcal{I}$ des polynômes $\neq 0$. L'ensemble de leurs degrés est une partie de $\mathbb{N}$, donc a un plus petit élément $n$. Soit $P \neq 0$ un polynôme de $\mathcal{I}$ ayant le degré $n$. Si $A \in \mathcal{I}$, on peut écrire $A = PQ + R$ avec $R = 0$ ou $0 \leqslant \deg R < n$. Comme $R = A - PQ \in \mathcal{I}$, la condition $0 \leqslant \deg R < n$ est impossible, donc $R = 0$ et $A = PQ$. Cela prouve que $\mathcal{I} \subset (P)$. Comme $(P) \subset \mathcal{I}$, on a l'égalité.
(ii) On a $(P) = (P')$ si et seulement si $P$ et $P'$ se divisent l'un l'autre, *i.e.* s'il existe $\lambda \in K, \lambda \neq 0$ tel que $P' = \lambda P$. □

**Définition 1. 19.** *Un idéal bilatère $\mathcal{I}$ d'un anneau $\mathcal{A}$ est dit maximal si $\mathcal{I} \neq \mathcal{A}$ et si tout idéal bilatère $\mathcal{J}$ tel que $\mathcal{I} \subset \mathcal{J} \subset \mathcal{A}$ vérifie $\mathcal{J} = \mathcal{I}$ ou $\mathcal{J} = \mathcal{A}$.*

**Exemples.**
(i) Un idéal $\mathcal{I} = (n) = n\mathbb{Z}$ de $\mathbb{Z}$ ($n \in \mathbb{N}$) est maximal si et seulement si $n$ est un nombre premier.
(ii) Un idéal $\mathcal{I} = (P)$ de $K[X]$ ($K$ corps commutatif) est maximal si et seulement si le polynôme $P$ est *irréductible*, *i.e.* $P \neq 0, \deg P \geqslant 1$, et les seuls diviseurs de $P$ sont les éléments $\lambda \in K$, $\lambda \neq 0$ ou les polynômes de la forme $\lambda P$ avec $\lambda \in K$, $\lambda \neq 0$.

### 1.14.4 Centre d'un anneau.

**Définition 1. 20.** *On appelle centre d'un anneau $\mathcal{A}$ le sous-ensemble*

$$Z(\mathcal{A}) = \{a \in \mathcal{A}\,; ax = xa \quad \forall x \in \mathcal{A}\}.$$

**Proposition 1. 11.** *Le centre $Z(\mathcal{A})$ d'un anneau $\mathcal{A}$ est un sous-anneau de $\mathcal{A}$.*

**Démonstration.**

On a $0 \cdot a = 0 = a \cdot 0$ pour tout $a \in \mathcal{A}$, donc $0 \in Z(\mathcal{A})$. Pour tous $x, y \in Z(\mathcal{A})$:
• $(x+y)a = xa + ya = ax + ay = a(x+y)$ $\forall a \in \mathcal{A}$, donc $x + y \in Z(\mathcal{A})$;
• $(-x)\,a = -(xa) = -(ax) = a\,(-x)$ $\forall a \in \mathcal{A}$, donc $-x \in Z(\mathcal{A})$;
• $xya = xay = axy$ $\forall a \in \mathcal{A}$, donc $xy \in Z(\mathcal{A})$. □

**Exemples.**

(i) Le centre de l'anneau $M_n(\mathbb{C})$ est l'ensemble des matrices de la forme $\lambda I$ où $\lambda \in \mathbb{C}$ et $I$ désigne la matrice identité.

(ii) Le centre du corps $\mathbb{H}$ des quaternions est $\left\{\begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix} \;; x \in \mathbb{R}\right\}$ qu'on peut identifier à $\mathbb{R}$ (voir ex. 1.20).

(iii) Le centre d'un anneau unitaire non commutatif n'est *jamais* un idéal bilatère. Sinon, comme il contient bien entendu 1, il contiendrait tout $a \in \mathcal{A}$ puisque $a = a \cdot 1$ donc ce serait $\mathcal{A}$ tout entier et $\mathcal{A}$ serait commutatif.

### 1.14.5 Anneau quotient.

Soit $\mathcal{A}$ un anneau et $\mathcal{I}$ un idéal bilatère de $\mathcal{A}$. On définit une relation $\mathcal{R}$ par $x\mathcal{R}y \Leftrightarrow x - y \in \mathcal{I}$. On voit facilement que $\mathcal{R}$ est une relation d'équivalence sur $\mathcal{A}$. On note au lieu de $x\mathcal{R}y$: $x \equiv y \pmod{\mathcal{I}}$. L'ensemble quotient est noté $\mathcal{A}/\mathcal{I}$.

**Théorème 1. 17.** *Soit $\mathcal{A}$ un anneau et $\mathcal{I}$ un idéal bilatère de $\mathcal{A}$.*
*(i) $\mathcal{A}/\mathcal{I}$ est un anneau. Si $\mathcal{A}$ est commutatif (resp. unitaire), $\mathcal{A}/\mathcal{I}$ est aussi commutatif (resp. unitaire).*
*(ii) On suppose $\mathcal{A}$ commutatif et unitaire. Alors $\mathcal{A}/\mathcal{I}$ est un corps si et seulement si l'idéal $\mathcal{I}$ est maximal.*

**Démonstration.**

(i) Le raisonnement est analogue à celui fait pour $\mathbb{Z}/n\mathbb{Z}$ (Th. 1.3). Il faut d'abord définir la somme et le produit de deux classes. Soient $[x], [y] \in \mathcal{A}/\mathcal{I}$. Les éléments $x, y \in \mathbb{Z}$ sont des *représentants* des classes respectives $[x], [y]$. Si $x', y'$ sont des autres représentants des classes respectives $[x], [y]$, on aura $x' \equiv x \pmod{\mathcal{I}}$, et $y' \equiv y \pmod{\mathcal{I}}$, donc $x' = x + h$ et $y' = y + k$ avec $h, k \in \mathcal{I}$. Alors $x' + y' = x + y + h + k$ donc $x' + y' \equiv x + y \pmod{\mathcal{I}}$ (puisque $h + k \in \mathcal{I}$) et $[x' + y'] = [x + y]$. Ainsi la classe $[x + y]$ *ne dépend pas des représentants $x, y$ utilisés pour* $[x], [y]$. On la note $[x] + [y]$. Cela définit une application $+ : \mathcal{A}/\mathcal{I} \times \mathcal{A}/\mathcal{I} \to \mathcal{A}/\mathcal{I}$. De même, $x'y' = (x + h)(y + k) = xy + hy + xk + hk$. Or comme $\mathcal{I}$ est un idéal bilatère, $hy \in \mathcal{I}, xk \in \mathcal{I}$ et $hk \in \mathcal{I}$, donc $hy + xk + hk \in \mathcal{I}$. Ainsi $x'y' \equiv xy \pmod{\mathcal{I}}$, $[x'y'] = [xy]$ et la classe $[xy]$ ne dépend pas des représentants $x, y$ utilisés pour $[x], [y]$. On la note $[x] \cdot [y]$ et cela définit une application $\cdot : \mathcal{A}/\mathcal{I} \times \mathcal{A}/\mathcal{I} \to \mathcal{A}/\mathcal{I}$. Les différents axiomes se vérifient alors sans difficulté.
(ii) Soit $\mathcal{I}$ un idéal de $\mathcal{A}$. L'anneau $\mathcal{A}/\mathcal{I}$ est un corps si et seulement si $[0] \neq [1]$, *i.e.* $\mathcal{I} \neq \mathcal{A}$, et tout élément non nul est inversible pour la multiplication, *i.e.*

$$\forall [x] \neq [0], \quad \exists [y] \qquad [x][y] = [1].$$

Comme $[x][y] = [xy]$, cette dernière condition équivaut à

$$\forall x \notin \mathcal{I}, \exists y \in \mathcal{A}, \exists h \in \mathcal{I} \qquad xy + h = 1. \tag{1.15}$$

Or on vérifie facilement que pour tout $x \in \mathcal{A}$, l'ensemble

$$(x) + \mathcal{I} = \{z \in \mathcal{A};\ \exists y \in \mathcal{A},\ \exists h \in \mathcal{I} \quad z = xy + h\}$$

est un idéal de $\mathcal{A}$. La condition (1.15) s'écrit simplement

$$\forall x \notin \mathcal{I}, \quad 1 \in (x) + \mathcal{I},$$

ou encore

$$\forall x \notin \mathcal{I}, \quad (x) + \mathcal{I} = \mathcal{A}. \tag{1.16}$$

Notons que si $x \notin \mathcal{I}$,

$$(x) + \mathcal{I} \underset{\neq}{\supset} \mathcal{I}. \tag{1.17}$$

Supposons maintenant $\mathcal{I}$ maximal. On a $\mathcal{I} \neq \mathcal{A}$. Soit $x \notin \mathcal{I}$. $\mathcal{I}$ étant maximal, on a d'après (1.17) $(x) + \mathcal{I} = \mathcal{A}$. Donc (1.16) est satisfaite et $\mathcal{A}/\mathcal{I}$ est un corps.

Réciproquement, supposons que $\mathcal{A}/\mathcal{I}$ soit un corps. Comme $[0] \neq [1]$, on a $\mathcal{I} \neq \mathcal{A}$. Soit $\mathcal{J}$ un idéal tel que $\mathcal{I} \subset \mathcal{J} \subset \mathcal{A}$. Si $\mathcal{J} \neq \mathcal{I}$, il existe $x \in \mathcal{J} \setminus \mathcal{I}$. Alors, comme $\mathcal{A}/\mathcal{I}$ est un corps, la condition (1.16) est vérifiée donc $(x) + \mathcal{I} = \mathcal{A}$. Or $x \in \mathcal{J}$ et $\mathcal{I} \subset \mathcal{J}$, donc $(x) + \mathcal{I} \subset \mathcal{J}$ et ainsi $\mathcal{J} = \mathcal{A}$. Cela prouve que $\mathcal{I}$ est maximal. □

### 1.14.6 Homomorphismes d'anneaux.

**Définition 1. 21.** *Soient $\mathcal{A}, \mathcal{A}'$ deux anneaux. Une application $f : \mathcal{A} \longrightarrow \mathcal{A}'$ est appelée un homorphisme d'anneaux si elle possède les deux propriétés suivantes.*

*(i) $f$ est un homomorphisme de groupes additifs.*

*(ii) $f(xy) = f(x)f(y) \quad \forall \quad x, y \in \mathcal{A}$.*

*Si en plus $f$ est bijectif, on dit que $f$ est un isomorphisme.*

**Remarque.** Soit $f : \mathcal{A} \longrightarrow \mathcal{A}'$ un homorphisme (*resp.* isomorphisme) d'anneaux. Si $\mathcal{A}$ et $\mathcal{A}'$ sont des corps et $f(1) = 1$, on dit que $f$ est un homomorphisme (*resp.* isomorphisme) de corps.

**Théorème 1. 18.** *Soit $f : \mathcal{A} \longrightarrow \mathcal{A}'$ un homomorphisme d'anneaux.*
*(i) $Ker f = f^{-1}(\{0\})$ est un idéal bilatère de $\mathcal{A}$. $f$ est injectif si et seulement si $Ker f = \{0\}$.*
*(ii) $Im\ f = f(\mathcal{A})$ est un sous-anneau de $\mathcal{A}'$.*

**Démonstration.**

(i) Comme $f$ est en particulier un homorphisme de groupes additifs, le Th. 1.6 montre que $\operatorname{Ker} f$ est un sous-groupe additif de $\mathcal{A}$ et que $f$ est injectif si et seulement si $\operatorname{Ker} f = \{0\}$. Maintenant, pour tout $x \in \operatorname{Ker} f$ et tout $a \in \mathcal{A}$, $f(ax) = f(a)f(x) = f(a) \cdot 0 = 0$ et $f(xa) = f(x)f(a) = 0 \cdot f(a) = 0$ donc $ax \in \operatorname{Ker} f$ et $xa \in \operatorname{Ker} f$. Ainsi $\operatorname{Ker} f$ est un idéal bilatère.

(ii) D'après le même théorème, $\operatorname{Im} f$ est un sous-groupe additif de $\mathcal{A}'$. Pour $x', y' \in \operatorname{Im} f$, il existe $x, y \in \mathcal{A}$ tels que $x' = f(x)$ et $y' = f(y)$, et l'on a

alors $x'y' = f(x)f(y) = f(xy)$, donc $x'y' \in \text{Im } f$. Im $f$ est ainsi stable pour la multiplication. Donc c'est un sous-anneau. □

**Définition 1. 22.** *Soit $f : \mathcal{A} \longrightarrow \mathcal{A}'$ un homomorphisme d'anneaux. L'idéal bilatère $Ker\, f = f^{-1}(\{0\})$ de $\mathcal{A}$ s'appelle le noyau de $f$, et le sous-anneau $Im\ f = f(\mathcal{A})$ de $\mathcal{A}'$ s'appelle l'image de $f$.*

**Théorème 1. 19.** *Soit $K, K'$ deux corps et $f : K \longrightarrow K'$ un homomorphisme de corps. Alors $f$ est injectif et $Im\ f$ est un sous-corps de $K'$ isomorphe à $K$.*

**Démonstration.**

Ker $f$ est un idéal bilatère du corps $K$ et Ker $f \neq K$ puisque $f(1) = 1$. Donc Ker $f = \{0\}$ car (Prop. 1.9) les seuls idéaux bilatères d'un corps $K$ sont $\{0\}$ et $K$. Ainsi $f$ est injectif. D'après le Th. 1.18, Im $f$ est un sous-anneau de $K'$. Montrons que c'est un sous-corps. On a $1 = f(1) \in \text{Im } f$. Il reste donc à voir que $\forall y \in \text{Im } f$ tel que $y \neq 0$, on a $y^{-1} \in \text{Im } f$. Soit donc $y \in \text{Im } f$, $y \neq 0$. Il existe $x \in K$ tel que $y = f(x)$, et $x \neq 0$ puisque $y \neq 0$. Alors $f(x^{-1})f(x) = f(x^{-1}x) = f(1) = 1$ d'où $y^{-1} = (f(x))^{-1} = f(x^{-1}) \in \text{Im } f$. Donc Im $f$ est un sous-corps de $K'$. Il est alors immédiat que $f$ est un isomorphisme de $K$ sur le corps Im $f$. □

### 1.14.7 Décomposition canonique d'un homomorphisme d'anneaux.

**Théorème 1. 20.** *Soit $f : \mathcal{A} \longrightarrow \mathcal{A}'$ un homomorphisme d'anneaux et $\pi : \mathcal{A} \longrightarrow \mathcal{A}/Ker\, f$ la projection canonique. Il existe un homomorphisme unique*

$$\tilde{f} : \mathcal{A}/Ker\, f \longrightarrow \mathcal{A}'$$

*rendant le diagramme suivant commutatif:*

$$\begin{array}{ccc} \mathcal{A} & \xrightarrow{f} & \mathcal{A}' \\ \pi \big\downarrow & \nearrow & \tilde{f} \\ \mathcal{A}/Ker\, f & & \end{array}$$

*L'homomorphisme $\tilde{f}$ est injectif et c'est un isomorphisme de $\mathcal{A}/Ker\, f$ sur $Im\ f$.*

**Démonstration.**

La décomposition de $f$ en tant qu'homomorphisme de groupes additifs montre qu'il existe un unique homomorphisme $\tilde{f}$ du groupe additif $\mathcal{A}/\text{Ker } f$ dans le groupe additif $\mathcal{A}'$ rendant le diagramme commutatif. Il reste donc simplement à vérifier que $\tilde{f}$ est un homomorphisme d'anneaux, *i.e.* $\tilde{f}([x][y]) = \tilde{f}([x])\tilde{f}([y])\ \forall x, y \in \mathcal{A}$. Or par définition de $\tilde{f}$, cela s'écrit $f(xy) = f(x)f(y)\ \forall x, y \in \mathcal{A}$ et cela signifie précisément que $f$ est un homomorphisme d'anneaux, ce qui est par hypothèse. □

### 1.14.8 Anneau produit direct.

**Théorème 1. 21.** *Soient $\mathcal{A}_1$ et $\mathcal{A}_2$ deux anneaux. L'ensemble produit cartésien $\mathcal{A}_1 \times \mathcal{A}_2 = \{(x, y); x \in \mathcal{A}_1, y \in \mathcal{A}_2\}$ muni des deux lois*

$$(x, y) + (x', y') = (x + x', y + y') \quad (1.18)$$

$$(x, y) \cdot (x', y') = (xx', yy') \quad (1.19)$$

*est un anneau. L'anneau $\mathcal{A}_1 \times \mathcal{A}_2$ est commutatif (resp. unitaire) si et seulement si $\mathcal{A}_1$ et $\mathcal{A}_2$ sont commutatifs (resp. unitaires).*

**Démonstration.**

La démonstration est analogue à celle du Th.1.14, et nous l'omettons. □

**Définition 1. 23.** *L'anneau $\mathcal{A}_1 \times \mathcal{A}_2$ est appelé l'anneau produit direct des deux anneaux $\mathcal{A}_1$ et $\mathcal{A}_2$.*

**Théorème 1. 22.** *Soient $p, q$ deux entiers $> 0$ premiers entre eux : $p \curlywedge q = 1$. Alors l'anneau $\mathbb{Z}/pq\mathbb{Z}$ est isomorphe à l'anneau produit direct $\mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}$.*

**Démonstration.**

La démonstration est analogue à celle du Th.1.15. L'application $f : \mathbb{Z} \to \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}$ définie par $f(x) = ([x]_p, [x]_q)$, où $[x]_p$ et $[x]_q$ désignent respectivement les classes de $x$ modulo $p$ et modulo $q$ est un homorphisme d'anneaux. Son noyau est $\operatorname{Ker} f = pq\mathbb{Z}$. La décomposition canonique de l'homomorphisme $f$ donne alors un homomorphisme injectif

$$\tilde{f} : \mathbb{Z}/pq\mathbb{Z} \longrightarrow \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}. \tag{1.20}$$

Comme $\mathbb{Z}/pq\mathbb{Z}$ et $\mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}$ ont le même cardinal $pq$, $\tilde{f}$ est bijectif, donc c'est un isomorphisme. □

## 1.15 Appendice : PGCD et PPCM dans $\mathbb{Z}$.

### 1.15.1 PGCD.

Soient $a, b \in \mathbb{Z}$. Considérons $a\mathbb{Z} + b\mathbb{Z}$. C'est un idéal de $\mathbb{Z}$. Donc il existe $D \in \mathbb{N}$ unique tel que

$$a\mathbb{Z} + b\mathbb{Z} = D\mathbb{Z}. \tag{1.21}$$

**Définition 1. 24.** *Le nombre $D \in \mathbb{N}$ défini par (1.21) est appelé plus grand commun diviseur (PGCD) de $a$ et $b$ et est noté $D = a \curlywedge b$. On dit que $a$ et $b$ sont premiers entre eux si $a \curlywedge b = 1$.*

**Proposition 1. 12.** *Soient $a, b \in \mathbb{Z}$ et $D = a \curlywedge b$.*
*(i) $D$ est le seul élément $x$ de $\mathbb{N}$ ayant la propriété suivante :*

*les diviseurs dans $\mathbb{Z}$ communs à $a$ et $b$ sont les diviseurs dans $\mathbb{Z}$ de $x$.*

*(ii) Il existe $u, v \in \mathbb{Z}$ tels que*

$$D = au + bv.$$

*(iii) $a$ et $b$ sont premiers entre eux si et seulement si il existe $u, v \in \mathbb{Z}$ tels que*

$$1 = au + bv \quad \textit{(identité de Bezout)} .$$

**Démonstration.**

(i) Si $p \in \mathbb{Z}$ divise $a$ et $b$, il divise tous les éléments de $a\mathbb{Z}+b\mathbb{Z}$, et réciproquement. Donc $p$ divise $a$ et $b$ si et seulement s'il divise tous les éléments de $D\mathbb{Z}$. Pour cela, il faut et il suffit que $p$ divise $D$. $D$ possède donc la propriété indiquée. Maintenant, soit $x$ un autre élément de $\mathbb{N}$ ayant cette propriété. $x$ divise $a$ et $b$, donc il divise $D$. Cela signifie que $D \in x\mathbb{Z}$ et implique donc $D\mathbb{Z} \subset x\mathbb{Z}$. De même, $D$ divise $a$ et $b$, donc il divise $x$, ce qui signifie $x \in D\mathbb{Z}$ et implique $x\mathbb{Z} \subset D\mathbb{Z}$. On en déduit que $D\mathbb{Z} = x\mathbb{Z}$. D'où $D = x$ puisque $D, x \in \mathbb{N}$.
(ii) On a par définition $D \in a\mathbb{Z} + b\mathbb{Z}$.
(iii) On a :

$$\begin{aligned} D = 1 &\Leftrightarrow a\mathbb{Z} + b\mathbb{Z} = \mathbb{Z} \\ &\Leftrightarrow 1 \in a\mathbb{Z} + b\mathbb{Z} \\ &\Leftrightarrow \exists u, v \in \mathbb{Z} \quad 1 = au + bv \end{aligned}$$

□

**Proposition 1. 13.** *Soient $a, b \in \mathbb{Z}$ et $p \in \mathbb{N}$. Alors*

$$(pa) \curlywedge (pb) = p(a \curlywedge b).$$

**Démonstration.**

$pa\mathbb{Z} + pb\mathbb{Z} = p(a\mathbb{Z} + b\mathbb{Z}) = p(a \curlywedge b)\mathbb{Z}$ donc $(pa) \curlywedge (pb) = p(a \curlywedge b)$ puisque $p, a \curlywedge b \in \mathbb{N}$. □

**Corollaire.** *Soient $a, b \in \mathbb{Z}$ et $d \in \mathbb{N}^*$ un diviseur commun à $a$ et $b$ :*

$$\begin{aligned} a &= da_1 \quad (a_1 \in \mathbb{Z}) \\ b &= db_1 \quad (b_1 \in \mathbb{Z}). \end{aligned}$$

*Alors $d = a \curlywedge b$ si et seulement si $a_1$ et $b_1$ sont premiers entre eux.*

**Démonstration.**

On a $a \curlywedge b = (da_1) \curlywedge (db_1) = d(a_1 \curlywedge b_1)$. Donc

$$d = a \curlywedge b \Leftrightarrow d = d(a_1 \curlywedge b_1) \Leftrightarrow a_1 \curlywedge b_1 = 1$$

puisque $d \neq 0$. □

**Proposition 1. 14.** *Soient $a, b, c \in \mathbb{Z}$. Si $a$ est premier avec $b$ et avec $c$, il est premier avec $bc$.*

**Démonstration.**

Il existe $u, v, u', v' \in \mathbb{Z}$ tels que $au + bv = 1$ et $au' + cv' = 1$. Par multiplication, on alors $a^2uu' + acuv' + bavu' + bcvv' = 1$, qui s'écrit $aw + bcz = 1$ avec $w = auu' + cuv' + bvu'$ et $z = vv'$. Comme $w, z \in \mathbb{Z}$, on obtient d'après l'identité de Bezout $a \curlywedge bc = 1$. □

**Corollaire.** *(i) Soient $a, b_1, \cdots, b_n \in \mathbb{Z}$. Si $a$ est premier avec chaque $b_i$, $a$ est premier avec le produit $b_1 \cdots b_n$.*
*(ii) Soient $a, b \in \mathbb{Z}$. Si $a$ est premier avec $b$, $a^k$ est premier avec $b^\ell$ pour tous $k, \ell \in \mathbb{N}^*$.*

**Démonstration.**

(i) $a$ est premier avec $b_1$ et $b_2$, donc avec $b_1b_2$. Etant premier avec $b_1b_2$ et $b_3$, il est premier avec $b_1b_2b_3$. Par récurrence, $a$ est premier avec $b_1 \cdots b_n$.
(ii) On obtient successivement par (i) $a \curlywedge b^\ell = 1$ et $a^k \curlywedge b^\ell = 1$. □

**Théorème 1. 23 (Théorème de Gauss).** *Soient $a, b, c \in \mathbb{Z}$. Si $a$ divise $bc$ et est premier avec $b$, alors $a$ divise $c$.*

**Démonstration.**

Comme $a$ est premier avec $b$, il existe $u, v \in \mathbb{Z}$ tels que $au+bv = 1$. Par multiplication par $c$ on obtient $acu + bcv = c$. Or $a$ divise $bc$, donc il divise $acu + bcv = c$. □

**Proposition 1. 15.** *Si $a \in \mathbb{Z}$ est divisible par deux entiers $b, c \in \mathbb{Z}$ premiers entre eux, $a$ est divisible par $bc$.*

**Démonstration.**

Il existe $u, v \in \mathbb{Z}$ tels que $bu + cv = 1$. Alors $a = abu + acv$. Si $q, q' \in \mathbb{Z}$ sont tels que $a = bq = cq'$ on aura donc $a = cq'bu + bqcv = bc(q'u + qv)$. □

## 1.15.2 Algorithme d'Euclide.

Soient $a, b \in \mathbb{N}^*$. Par division euclidienne, il existe $q_1, r_1$ tels que $a = bq_1 + r_1$ avec $q_1, r_1 \in \mathbb{N}$, $0 \leqslant r_1 < b$. Si $r_1 = 0$, un entier $p$ divise $a$ et $b$ si et seulement si il divise $b$, donc dans ce cas $a \curlywedge b = b$. Si $r_1 \neq 0$, $p$ divise $a$ et $b$ si et seulement si il divise $b$ et $r_1$, donc dans ce cas $a \curlywedge b = b \curlywedge r_1$. On applique alors le même raisonnement au couple $(b, r_1)$ : $b = r_1q_2 + r_2$ avec $q_2, r_2 \in \mathbb{N}$, $0 \leqslant r_2 < b$. Si $r_2 = 0$, $b \curlywedge r_1 = r_1$ donc $a \curlywedge b = r_1$. Si $r_2 \neq 0$, $b \curlywedge r_1 = r_1 \curlywedge r_2$ donc $a \curlywedge b = r_1 \curlywedge r_2$. On est alors amené à effectuer les divisions suivantes en s'arrêtant dès que l'on obtient un reste nul.

$$\begin{aligned}
a &= bq_1 + r_1 \quad \text{avec } q_1, r_1 \in \mathbb{N},\ 0 < r_1 < b,\\
b &= r_1q_2 + r_2 \quad \text{avec } q_2, r_2 \in \mathbb{N},\ 0 < r_2 < r_1,\\
\vdots &\ \ \vdots\ \ \vdots\\
r_n &= r_{n+1}q_{n+2} + r_{n+2} \text{ avec } q_{n+2}, r_{n+2} \in \mathbb{N},\ 0 < r_{n+2} < r_{n+1},\\
r_{n+1} &= r_{n+2}q_{n+3} \text{ avec } q_{n+3} \in \mathbb{N}.
\end{aligned}$$

Alors $a \curlywedge b = b \curlywedge r_1 = r_1 \curlywedge r_2 = \cdots = r_{n+1} \curlywedge r_{n+2} = r_{n+2}$. Le PGCD $a \curlywedge b$ de $a$ et $b$ est donc le dernier reste non nul. C'est l'algorithme d'Euclide pour la recherche du PGCD.

## 1.15.3 PPCM.

Soient $a, b \in \mathbb{Z}$. Considérons $a\mathbb{Z} \cap b\mathbb{Z}$. C'est un idéal de $\mathbb{Z}$. Donc il existe $M \in \mathbb{N}$ unique tel que

$$a\mathbb{Z} \cap b\mathbb{Z} = M\mathbb{Z}. \tag{1.22}$$

**Définition 1. 25.** *Le nombre $M \in \mathbb{N}$ défini par (1.22) est appelé plus petit commun multiple (PPCM) de $a$ et $b$ et est noté $M = a \curlyvee b$.*

**Proposition 1. 16.** *Soient $a, b \in \mathbb{Z}$ et $M = a \curlyvee b$. $M$ est le seul élément $x$ de $\mathbb{N}$ ayant la propriété suivante :*

*les multiples dans $\mathbb{Z}$ communs à $a$ et $b$ sont les multiples dans $\mathbb{Z}$ de $x$.*

**Démonstration.**

Soit $x \in \mathbb{N}$ quelconque. $a\mathbb{Z} \cap b\mathbb{Z}$ est l'ensemble des multiples communs dans $\mathbb{Z}$ à $a$ et $b$, et $x\mathbb{Z}$ est l'ensemble des multiples dans $\mathbb{Z}$ de $x$. Donc $x$ vérifie la condition de l'énoncé si et seulement si $x\mathbb{Z} = a\mathbb{Z} \cap b\mathbb{Z}$. Or $a\mathbb{Z} \cap b\mathbb{Z} = M\mathbb{Z}$. Donc $x$ vérifie la condition de l'énoncé si et seulement si $x\mathbb{Z} = M\mathbb{Z}$, *i.e.* $x = M$ puisque $x, M \in \mathbb{N}$. □

**Proposition 1. 17.** *Soient $a, b \in \mathbb{Z}$, $D = a \curlywedge b$ leur PGCD et $M = a \curlyvee b$ leur PPCM. Alors $|ab| = DM$.*

**Démonstration.**

On peut supposer $a, b \in \mathbb{N}$. Si $a$ ou $b$ est nul , on a $M = 0$ et l'équation est triviale. On suppose donc $a, b \in \mathbb{N}^*$. Cela implique $D \neq 0$. On a $a = Da_1$ et $b = Db_1$ avec $a_1$ et $b_1$ premiers entre eux d'après le Corollaire de la Prop. 1.13. Alors $ab = D(Da_1b_1)$. Il faut donc démontrer que $Da_1b_1 = M$. Cela équivaut à montrer que les multiples communs à $a$ et $b$ sont les multiples de $Da_1b_1$ (Prop.1.16). $Da_1b_1$ est multiple de $a$ et de $b$, donc tout multiple de $Da_1b_1$ aussi. Réciproquement, soit $p \in \mathbb{Z}$ un multiple commun à $a$ et $b$. Il existe $q \in \mathbb{Z}$ tel que $p = aq$, ce qui donne $p = Da_1q$. De même, il existe $s \in \mathbb{Z}$ tel que $p = bs$, ce qui donne $p = Db_1s$. Ainsi $Da_1q = Db_1s$ d'où $a_1q = b_1s$ puisque $D \neq 0$. Maintenant, $b_1$ divise $a_1q$ et est premier avec $a_1$, donc divise $q$ d'après le Théorème de Gauss. Soit $t \in \mathbb{Z}$ tel que $q = b_1t$. On a alors $p = Da_1q = Da_1b_1t$. Donc $p$ est multiple de $Da_1b_1$. □

### 1.15.4 Cas d'une famille finie d'éléments de $\mathbb{Z}$.

Les notions de PGCD et PPCM se généralisent au cas d'une famille finie d'éléments $a_1, \cdots, a_k$ ($k \geqslant 2$) de $\mathbb{Z}$. En effet,

$$a_1\mathbb{Z} + \cdots + a_k\mathbb{Z} = \{a_1p_1 + \cdots + a_kp_k \ ; \ p_1, \cdots, p_k \in \mathbb{Z}\}$$

est un idéal de $\mathbb{Z}$, donc il existe $D \in \mathbb{N}$ unique tel que

$$a_1\mathbb{Z} + \cdots + a_k\mathbb{Z} = D\mathbb{Z}. \tag{1.23}$$

Le nombre $D \in \mathbb{N}$ défini par (1.23) est appelé plus grand commun diviseur (PGCD) de la famille $a_1, \cdots, a_k$. On dit que $a_1, \cdots, a_k$ sont *premiers entre eux dans leur ensemble* si $D = 1$. De même,

$$a_1\mathbb{Z} \cap \cdots \cap a_k\mathbb{Z}$$

est un idéal de $\mathbb{Z}$, donc il existe $M \in \mathbb{N}$ unique tel que

$$a_1\mathbb{Z} \cap \cdots \cap a_k\mathbb{Z} = M\mathbb{Z}. \tag{1.24}$$

Le nombre $M \in \mathbb{N}$ défini par (1.24) est appelé plus petit commun multiple (PPCM) de la famille $a_1, \cdots, a_k$.

Notons que $(a,b) \to a \curlywedge b$ est une loi interne *associative* sur $\mathbb{Z}$. On a en effet pour $a, b, c \in \mathbb{Z}$

$$\begin{aligned} ((a \curlywedge b) \curlywedge c)\mathbb{Z} &= (a \curlywedge b)\mathbb{Z} + c\mathbb{Z} \\ &= (a\mathbb{Z} + b\mathbb{Z}) + c\mathbb{Z} \\ &= a\mathbb{Z} + b\mathbb{Z} + c\mathbb{Z} \\ &= a\mathbb{Z} + (b\mathbb{Z} + c\mathbb{Z}) \\ &= a\mathbb{Z} + (b \curlywedge c)\mathbb{Z} \\ &= (a \curlywedge (b \curlywedge c))\mathbb{Z} \end{aligned}$$

donc

$$(a \curlywedge b) \curlywedge c = a \curlywedge (b \curlywedge c).$$

D'après le Lemme 1.1 valable pour toute loi associative, on a par récurrence

$$a_1\mathbb{Z} + \cdots + a_k\mathbb{Z} = (a_1 \curlywedge \cdots \curlywedge a_k)\mathbb{Z}$$

donc le PGCD de la famille $a_1, \cdots, a_k$ est $a_1 \curlywedge \cdots \curlywedge a_k$.

De même, $(a,b) \to a \curlyvee b$ est une loi interne associative sur $\mathbb{Z}$, car pour $a, b, c \in \mathbb{Z}$

$$\begin{aligned} ((a \curlyvee b) \curlyvee c)\mathbb{Z} &= (a \curlyvee b)\mathbb{Z} \cap c\mathbb{Z} \\ &= (a\mathbb{Z} \cap b\mathbb{Z}) \cap c\mathbb{Z} \\ &= a\mathbb{Z} \cap b\mathbb{Z} \cap c\mathbb{Z} \\ &= a\mathbb{Z} \cap (b\mathbb{Z} \cap c\mathbb{Z}) \\ &= a\mathbb{Z} \cap (b \curlyvee c)\mathbb{Z} \\ &= (a \curlyvee (b \curlyvee c))\mathbb{Z} \end{aligned}$$

donc

$$(a \curlyvee b) \curlyvee c = a \curlyvee (b \curlyvee c).$$

On a par récurrence

$$a_1\mathbb{Z} \cap \cdots \cap a_k\mathbb{Z} = (a_1 \curlyvee \cdots \curlyvee a_k)\mathbb{Z}$$

donc le PPCM de la famille $a_1, \cdots, a_k$ est $a_1 \curlyvee \cdots \curlyvee a_k$.

Les propositions 1.12 et 1.16 se généralisent immédiatement sous la forme suivante.

**Proposition 1. 18.** *Soient* $a_1, \cdots, a_k \in \mathbb{Z}$ $(k \geqslant 2)$ *et* $D = a_1 \curlywedge \cdots \curlywedge a_k$.
*(i) D est le seul élément x de* $\mathbb{N}$ *ayant la propriété suivante :*

*les diviseurs dans* $\mathbb{Z}$ *communs à* $a_1, \cdots, a_k$ *sont les diviseurs dans* $\mathbb{Z}$ *de x.*

*(ii) Il existe* $u_1, \cdots, u_k \in \mathbb{Z}$ *tels que*

$$D = a_1u_1 + \cdots + a_ku_k.$$

*(iii)* $a_1, \cdots, a_k$ *sont premiers entre eux dans leur ensemble si et seulement si il existe* $u_1, \cdots, u_k \in \mathbb{Z}$ *tels que*

$$1 = a_1u_1 + \cdots + a_ku_k \quad \textit{(identité de Bezout)}.$$

**Proposition 1. 19.** *Soient* $a_1, \cdots, a_k \in \mathbb{Z}$ $(k \geqslant 2)$ *et* $M = a_1 \curlyvee \cdots \curlyvee a_k$. $M$ *est le seul élément* $x$ *de* $\mathbb{N}$ *ayant la propriété suivante :*

*les multiples dans* $\mathbb{Z}$ *communs à* $a_1, \cdots, a_k$ *sont les multiples dans* $\mathbb{Z}$ *de* $x$.

La Proposition 1.13 se généralise par récurrence; son Corollaire se généralise en :

**Corollaire.** *Soient* $a_1, \cdots, a_k \in \mathbb{Z}$ $(k \geqslant 2)$ *et* $d \in \mathbb{N}^*$ *un diviseur commun à* $a_1, \cdots, a_k$ :

$$a_i = d(a_i)_1 \quad ((a_i)_1 \in \mathbb{Z}) \quad \forall i\; 1 \leqslant i \leqslant k\,.$$

*Alors* $d = a_1 \curlywedge \cdots \curlywedge a_k$ *si et seulement si* $(a_1)_1, \cdots, (a_k)_1$ *sont premiers entre eux dans leur ensemble.*

On notera que la Proposition 1.17 ne se généralise pas pour $k > 2$. Par exemple, on a

$$2 \cdot 4 \cdot 6 = 48 \neq (2 \curlywedge 4 \curlywedge 6)(2 \curlyvee 4 \curlyvee 6) = 2 \cdot 12 = 24.$$

### 1.15.5 Décomposition en facteurs premiers.

**Définition 1. 26.** *Un nombre* $p \in \mathbb{N}^*$ *est dit premier si* $p \neq 1$ *et les seuls diviseurs de* $p$ *dans* $\mathbb{N}$ *sont* 1 *et* $p$.

**Lemme 1. 6.** *Soit* $a \in \mathbb{N}^*$, $a > 1$. *Il existe un nombre premier* $p$ *qui est un diviseur de* $a$.

**Démonstration.**
Si $a$ est premier, c'est trivial. Supposons donc $a$ non premier. Il existe un diviseur $p_1 \in \mathbb{N}^*$ de $a$ tel que $1 < p_1 < a$. Si $p_1$ est premier, on prend $p = p_1$. Sinon, il existe de même un diviseur $p_2 \in \mathbb{N}^*$ de $p_1$ tel que $1 < p_2 < p_1$. Si $p_2$ est premier, on prend $p = p_2$. Sinon, on itère le processus. Il ne peut pas continuer indéfiniment puisque la suite $p_1, p_2, \cdots$ est une suite d'entiers $> 1$ strictement décroissante. Au bout d'un nombre fini $k$ d'étapes on obtiendra donc $p_k$ premier tel que $1 < p_k < p_{k-1} < \cdots < p_2 < p_1 < a$ et que $p_k$ divise $p_{k-1}$, $p_{k-1}$ divise $p_{k-2}$, ..., $p_2$ divise $p_1$, $p_1$ divise $n$. Alors $p = p_k$ convient. □

**Théorème 1. 24.** *Soit* $a \in \mathbb{Z}^*$, $a \neq \pm 1$. *Il existe une décomposition unique*

$$a = \pm\, (p_1)^{\alpha_1} \cdots (p_r)^{\alpha_r} \tag{1.25}$$

*où* $r \in \mathbb{N}^*$, $p_1, \cdots, p_r$ *sont des nombres premiers vérifiant* $p_1 < p_2 < \cdots < p_r$ *et* $\alpha_1, \cdots, \alpha_r \in \mathbb{N}^*$.

**Démonstration.**
On peut supposer $a > 0$. Si $a$ est premier, l'assertion du Théorème est triviale, avec $r = 1, p_1 = a, \alpha_1 = 1$. On suppose donc $a$ non premier.
*Existence.* D'après le Lemme 1.6, il existe un nombre premier $s_1$ qui est un diviseur de $a$, *i.e.* $a = s_1 q_1$ avec $q_1 \in \mathbb{N}^*$, $1 < q_1 < a$. Si $q_1$ est premier, on pose $s_2 = q_1$, et l'on a $a = s_1 s_2$. Si $q_1$ n'est pas premier, il existe un diviseur premier $s_2$ de $q_1$, *i.e.* $q_1 = s_2 q_2$ avec $q_2 \in \mathbb{N}^*$, $1 < q_2 < q_1$. Si $q_2$ est premier, on pose $s_3 = q_2$, et l'on a $a = s_1 s_2 s_3$. Sinon, on itère le processus. Il ne peut pas continuer indéfiniment puisque la suite $q_1, q_2, \cdots$ est une suite d'entiers $> 1$ strictement décroissante. Au bout d'un nombre fini $t$ d'étapes on obtiendra donc $q_t$ premier et alors, en posant

$s_{t+1} = q_t$, $a = s_1 \cdots s_{t+1}$ avec $s_1, \cdots, s_t$ premiers. Maintenant, il suffit de mettre $s_1, \cdots, s_{t+1}$ dans l'ordre croissant et de regrouper ceux qui sont égaux pour obtenir une décomposition du type (1.25).

*Unicité.* Soient 2 décompositions

$$a = (p_1)^{\alpha_1} \cdots (p_r)^{\alpha_r} = (q_1)^{\beta_1} \cdots (q_s)^{\beta_s}$$

avec $r, s \in \mathbb{N}^*$, $p_1, \cdots, p_r$, $q_1, \cdots, q_s$ des nombres premiers vérifiant $p_1 < p_2 < \cdots < p_r$, $q_1 < q_2 < \cdots < q_s$ et $\alpha_1, \cdots, \alpha_r, \beta_1, \cdots, \beta_s \in \mathbb{N}^*$. Pour tout $i$ $(1 \leqslant i \leqslant r)$, il existe $j$ $(1 \leqslant j \leqslant s)$ tel que $p_i = q_j$. Sinon il existerait $i_0$ $(1 \leqslant i \leqslant r)$ tel que $p_{i_0} \neq q_j$ $\forall j$ $(1 \leqslant j \leqslant s)$ et alors $p_{i_0}$ serait premier avec tous les $q_j$ donc avec $(q_1)^{\beta_1} \cdots (q_s)^{\beta_s} = a$ (d'après le Corollaire du Th. de Gauss), ce qui est absurde. On obtient donc une application $\sigma$ de $\{1, \cdots, r\}$ dans $\{1, \cdots, s\}$ telle que $p_i = q_{\sigma(i)}$ $\forall i$. Elle est injective puisque $\sigma(i) = \sigma(i')$ implique $p_i = q_{\sigma(i)} = q_{\sigma(i')} = p_{i'}$ donc $i = i'$. Elle est aussi surjective car s'il existait un $j_0$ $(1 \leqslant j_0 \leqslant s)$ tel que $j_0 \neq \sigma(i)$ $\forall i$, $q_{j_0}$ serait premier avec tous les $q_{\sigma(i)} = p_i$ et diviserait $a$ ce qui est absurde. Donc $\sigma$ est bijective et $r = s$. Alors les conditions $p_1 = q_{\sigma(1)} < p_2 = q_{\sigma(2)} < \cdots < p_r = q_{\sigma(r)}$ et $q_1 < q_2 < \cdots < q_r$ impliquent $\sigma(1) = 1, \cdots, \sigma(r) = r$, donc $p_1 = q_1, \cdots, p_r = q_r$. Maintenant, pour tout $i$, $(p_i)^{\alpha_i} = (q_i)^{\alpha_i}$ divise $a$ et est premier avec $(q_m)^{\beta_m}$ pour tout $m \neq i$, donc avec $\prod_{m \neq i} (q_m)^{\beta_m}$. D'après le Th. de Gauss, cela implique que $(q_i)^{\alpha_i}$ divise $(q_i)^{\beta_i}$. Donc $\alpha_i \leqslant \beta_i$ $\forall i$ $(1 \leqslant i \leqslant r)$. De même, $\beta_i \leqslant \alpha_i$ $\forall i$ $(1 \leqslant i \leqslant r)$. Il vient donc $\alpha_i = \beta_i$ $\forall i$ $(1 \leqslant i \leqslant r)$. □

**Corollaire.** *Soient $a_1, \cdots, a_k \in \mathbb{Z}^*$ et $p_1 < \cdots < p_r$ les nombres premiers entrant dans la décomposition en facteurs premiers des $a_i$ différents de $\pm 1$. Soit pour chaque $i$, $(1 \leqslant i \leqslant k)$ $\alpha_1^i, \cdots, \alpha_r^i$ l'unique suite d'entiers $\geqslant 0$ tels que*

$$a_i = \pm (p_1)^{\alpha_1^i} \cdots (p_r)^{\alpha_r^i} = \pm \prod_{j=1}^{r} (p_j)^{\alpha_j^i}.$$

*(i) Les éléments de $\mathbb{Z}$ qui sont des diviseurs communs à $a_1, \cdots, a_k$ sont les éléments $b \in \mathbb{Z}$ de la forme*

$$b = \pm (p_1)^{\gamma_1} \cdots (p_r)^{\gamma_r} = \pm \prod_{j=1}^{r} (p_j)^{\gamma_j}.$$

*avec $0 \leqslant \gamma_j \leqslant \inf_{1 \leqslant i \leqslant k} \alpha_j^i$.*

*(ii)*

$$a_1 \curlywedge \cdots \curlywedge a_k = (p_1)^{\inf_{1 \leqslant i \leqslant k} \alpha_1^i} \cdots (p_r)^{\inf_{1 \leqslant i \leqslant k} \alpha_r^i} = \prod_{j=1}^{r} (p_j)^{\inf_{1 \leqslant i \leqslant k} \alpha_j^i}.$$

*(iii) Les éléments de $\mathbb{Z}$ qui sont des multiples communs à $a_1, \cdots, a_k$ sont les éléments $b \in \mathbb{Z}$ de la forme*

$$b = \pm (p_1)^{\lambda_1} \cdots (p_r)^{\lambda_r} = \pm \prod_{j=1}^{r} (p_j)^{\lambda_j}.$$

*avec $\lambda_j \geqslant \sup_{1 \leqslant i \leqslant k} \alpha_j^i$.*

*(iv)*

$$a_1 \curlyvee \cdots \curlyvee a_k = (p_1)^{\sup_{1 \leqslant i \leqslant k} \alpha_1^i} \cdots (p_r)^{\sup_{1 \leqslant i \leqslant k} \alpha_r^i} = \prod_{j=1}^{r} (p_j)^{\sup_{1 \leqslant i \leqslant k} \alpha_j^i}.$$

**Démonstration.**

La démonstration est immédiate et laissée au lecteur. □

## 1.16 Exercices.

**Exercice 1.1.**

Soit $x \in \mathbb{N}$ et $x = a_0 \cdot 1 + a_1 \cdot 10 + a_2 \cdot 10^2 + \ldots + a_p \cdot 10^p$ son écriture décimale ($0 \leqslant a_0, \ldots, a_p \leqslant 9$). Enoncer et démontrer les critères de divisibilité par : 3, 4, 5, 7, 8, 9, 11.

**Indication.**

• Divisibilité par 3. Dans le corps $\mathbb{Z}/3\mathbb{Z}$, on a $[10] = [1]$ donc $[10]^k = [1]$ $\forall k \in \mathbb{N}$ et

$$[x] = [a_0] + \cdots + [a_p] = [a_0 + \cdots + a_p].$$

Ainsi $x$ est divisible par 3 si et seulement si $a_0 + \cdots + a_p$ l'est.

• Divisibilité par 4. Dans l'anneau $\mathbb{Z}/4\mathbb{Z}$, on a $[10] = [2]$ et $[10]^k = [0]$ $\forall k \geqslant 2$. Donc

$$[x] = [a_0] + [a_1][2] = [a_0 + 2a_1].$$

Ainsi $x$ est divisible par 4 si et seulement si $a_0 + 2a_1$ l'est.

• Divisibilité par 5. Dans le corps $\mathbb{Z}/5\mathbb{Z}$, on a $[10]^k = [0]$ $\forall k \geqslant 1$. Donc

$$[x] = [a_0].$$

Ainsi $x$ est divisible par 5 si et seulement si $a_0$ l'est.

• Divisibilité par 7. Dans le corps $\mathbb{Z}/7\mathbb{Z}$, on a $[10] = [3]$, $[10]^2 = [3]^2 = [9] = [2]$, $[10]^3 = [3][2] = [6]$, $[10]^4 = [3][6] = [18] = [4]$, $[10]^5 = [3][4] = [12] = [5]$, $[10]^6 = [3][5] = [15] = [1]$. Alors pour tout $k \in \mathbb{N}$ $[10]^{6k} = [1]$, $[10]^{6k+1} = [3]$, $[10]^{6k+2} = [2]$, $[10]^{6k+3} = [6]$, $[10]^{6k+4} = [4]$, $[10]^{6k+5} = [5]$. Donc

$$\begin{aligned} [x] &= \sum_{k\geqslant 0}([a_{6k}] + [3][a_{6k+1}] + [2][a_{6k+2}] + [6][a_{6k+3}] + [4][a_{6k+4}] + [5][a_{6k+5}]) \\ &= \sum_{k\geqslant 0}([a_{6k}] + [3][a_{6k+1}] + [2][a_{6k+2}] - [a_{6k+3}] - [3][a_{6k+4}] - [2][a_{6k+5}]) \end{aligned}$$

puisque $[6] = -[1], [4] = -[3]$ et $[5] = -[2]$. Ainsi $x$ est divisible par 7 si et seulement si

$$\sum_{k\geqslant 0}([a_{6k}] + [3][a_{6k+1}] + [2][a_{6k+2}]) = \sum_{k\geqslant 0}([a_{6k+3}] + [3][a_{6k+4}] + [2][a_{6k+5}])$$

ou encore

$$\sum_{k\geqslant 0}(a_{6k} + 3a_{6k+1} + 2a_{6k+2}) \equiv \sum_{k\geqslant 0}(a_{6k+3} + 3a_{6k+4} + 2a_{6k+5}) \pmod 7.$$

Par exemple, pour un nombre de 6 chiffres, $x = \sum_{i=0}^{5} a_i \cdot 10^i$ , la condition s'écrit

$$a_0 + 3a_1 + 2a_2 \equiv a_3 + 3a_4 + 2a_5 \pmod 7.$$

Signalons aussi le critère suivant utilisé au Collège : si $N$ désigne *le nombre de dizaines* de $x$ (et non pas le chiffre des dizaines), *i.e.* $x = a_0 + N \cdot 10$, alors $[x] = [a_0] + [3][N]$ donc

$$[x] = [0] \Leftrightarrow [3][N] + [a_0] = [0] \Leftrightarrow [N] + [5][a_0] = [0] \Leftrightarrow [N] = [2][a_0] \quad .$$

car $[5][3] = [1]$ et $[5] = -[2]$. Ainsi $x$ est divisible par 7 si et seulement si $N - 2a_0$ l'est.
• Divisibilité par 8. Dans l'anneau $\mathbb{Z}/8\mathbb{Z}$, on a $[10] = [2]$, $[10]^2 = [4]$ et $[10]^k = [0]\ \forall k \geqslant 3$. Donc

$$[x] = [a_0] + [a_1][2] + [a_2][4] = [a_0 + 2a_1 + 4a_2].$$

Ainsi $x$ est divisible par 8 si et seulement si $a_0 + 2a_1 + 4a_2$ l'est.
• Divisibilité par 9. Dans l'anneau $\mathbb{Z}/9\mathbb{Z}$, on a $[10] = [1]$ donc $[10]^k = [1]$ $\forall k \in \mathbb{N}$ et

$$[x] = [a_0] + \cdots + [a_p] = [a_0 + \cdots + a_p].$$

Ainsi $x$ est divisible par 9 si et seulement si $a_0 + \cdots + a_p$ l'est.
• Divisibilité par 11. Dans le corps $\mathbb{Z}/11\mathbb{Z}$, on a $[10] = -[1]$, $[10]^2 = [1]$. Alors pour tout $k \in \mathbb{N}$ $[10]^{2k} = [1]$ et $[10]^{2k+1} = [10] = -[1]$. Donc

$$[x] = \sum_{k \geqslant 0} ([a_{2k}] - [a_{2k+1}]).$$

Ainsi $x$ est divisible par 11 si et seulement si

$$\sum_{k \geqslant 0} ([a_{2k}] - [a_{2k+1}]) = [0]$$

ou encore

$$\sum_{k \geqslant 0} a_{2k} \equiv \sum_{k \geqslant 0} a_{2k+1} \pmod{11}.$$

**Exercice 1.2.**

Montrer qu'un groupe $G$ tel que $x^2 = e \quad \forall x \in G$ est abélien.

**Exercice 1.3.**

Montrer qu'il n'y a, à isomorphisme près, que 2 groupes d'ordre 4, à savoir $\mathbb{Z}_4$ et $\mathbb{Z}_2 \times \mathbb{Z}_2$.

**Exercice 1.4.**

Montrer que si $p$ est un nombre premier, $\mathbb{Z}_p$ est à isomorphisme près le seul groupe d'ordre $p$,

**Indication.**

Soit $G$ un groupe d'ordre $p$ et $x$ un élément quelconque de $G$ différent de $e$. D'après le théorème de Lagrange, l'ordre de $\langle x \rangle$ est un diviseur de $p$. Or $p$ est premier, donc ses seuls diviseurs sont 1 ou $p$. Mais $x$ n'est pas d'ordre 1, puisque le seul élément d'ordre 1 est $e$ et que $x \neq e$. Donc l'ordre de $\langle x \rangle$ est $p$, et alors $\langle x \rangle = G$. $G$ est donc cyclique d'ordre $p$. D'après le Th. 1.10 on a $G \cong \mathbb{Z}_p$.

**Exercice 1.5.**

On considère le groupe symétrique $G = \mathcal{S}_n$, avec $n \geqslant 3$.

(i) Si $\tau = (a, b) \in \mathcal{S}_n$ est une transposition, calculer $s\tau s^{-1}$ pour $s \in \mathcal{S}_n$.

(ii) En déduire le centre $Z(\mathcal{S}_n)$.

**Indication.**

(i) $s(a, b)s^{-1} = (s(a), s(b))$.

(ii)Soit $s \in Z(\mathcal{S}_n)$. On va montrer que $s = Id$. Soit $a \in \{1, \ldots, n\}$. Comme $n \geqslant 3$, il existe $b, c \in \{1, \ldots, n\}$ tels que $a, b, c$ soient deux-à-deux distincts. Comme $s$ commute avec $(a, b)$, on a d'après (i) $(s(a), s(b)) = (a, b)$. Cela implique $\{s(a), s(b)\} = \{a, b\}$, donc $s(a) \in \{a, b\}$. De même, comme $s$ commute avec $(a, c)$, on a d'après (i) $(s(a), s(c)) = (a, c)$ et cela implique $\{s(a), s(c)\} = \{a, c\}$, donc

$s(a) \in \{a, c\}$. Ainsi $s(a) \in \{a, b\} \cap \{a, c\} = \{a\}$. Donc $s(a) = a$. Comme $a$ est arbitraire, $s$ est la permutation identité $Id$. On a donc $Z(\mathcal{S}_n) \subset \{Id\}$. Mais $Id$ appartient au centre, donc $Z(\mathcal{S}_n) = \{Id\}$.

**Exercice 1.6.**

Soit $U(\mathbb{Z} / n\mathbb{Z})$ l'ensemble des *unités de l'anneau* $\mathbb{Z} / n\mathbb{Z}$ , i.e. les éléments de l'anneau $\mathbb{Z} / n\mathbb{Z}$ qui sont inversibles pour la multiplication.

(i) Montrer que $[p] \in U(\mathbb{Z} / n\mathbb{Z}) \Leftrightarrow p \curlywedge n = 1$.

(ii) Montrer que $U(\mathbb{Z} / n\mathbb{Z})$ est un groupe abélien.

(iii) Déterminer le groupe $U(\mathbb{Z} / 8\mathbb{Z})$.

**Exercice 1.7.**

Montrer que le groupe $\mathrm{Aut}(\mathbb{Z}_n)$ est isomorphe au groupe $U(\mathbb{Z} / n\mathbb{Z})$ (voir ex. 1.6).

**Exercice 1.8.**

Soit $G$ un groupe.

(i) Montrer que $\mathrm{Int}(G) = \{\mathrm{Int}(x); x \in G\}$ est un sous-groupe distingué de $\mathrm{Aut}(G)$.

(ii) Montrer que $\mathrm{Int}(G)$ est isomorphe au groupe-quotient $G/ Z(G)$.

**Exercice 1.9.**

Soit $G$ l'ensemble $\mathbb{R} \times \mathbb{R}^*$ muni de la loi

$$(b, a)(b', a') = (b + ab', aa').$$

Montrer que $G$ est un groupe isomorphe au sous-groupe de $\mathrm{Bij}(\mathbb{R})$ dont les éléments sont les applications $x \mapsto ax + b, (b, a) \in \mathbb{R} \times \mathbb{R}^*$.

**Exercice 1.10.**

(i) A quelle condition nécessaire et suffisante sur $a, b > 0$ le sous-groupe additif $a\mathbb{Z} + b\mathbb{Z}$ est-il dense dans $\mathbb{R}$?

(ii) On suppose cette condition réalisée. Montrer que $a\mathbb{N} + b\mathbb{Z}$ est aussi dense dans $\mathbb{R}$.

**Indication.**

(i) Il est immédiat que $a\mathbb{Z} + b\mathbb{Z}$ est un sous-groupe de $\mathbb{R}$. Supposons que $a\mathbb{Z} + b\mathbb{Z}$ soit non dense. Alors il existe $c \geqslant 0$ tel que $a\mathbb{Z} + b\mathbb{Z} = c\mathbb{Z}$, et l'on a nécessairement $c > 0$ puisque $a\mathbb{Z} + b\mathbb{Z} \neq \{0\}$. $a \in a\mathbb{Z} + b\mathbb{Z}$, donc $a \in c\mathbb{Z}$. Il existe donc $p \in \mathbb{Z}^*$ tel que $a = cp$. De même, il existe $q \in \mathbb{Z}^*$ tel que $b = cq$. On a alors $\frac{a}{b} = \frac{p}{q} \in \mathbb{Q}$. Réciproquement, supposons $\frac{a}{b} \in \mathbb{Q}$. Il existe donc $p, q \in \mathbb{N}^*$ tels que $\frac{a}{b} = \frac{p}{q}$, et l'on peut supposer $p$ et $q$ premiers entre eux. On a $a = b\frac{p}{q}$ donc $a\mathbb{Z} + b\mathbb{Z} = \frac{bp}{q}\mathbb{Z} + b\mathbb{Z} = \frac{b}{q}(p\mathbb{Z} + q\mathbb{Z})$. Or $p\mathbb{Z} + q\mathbb{Z} = \mathbb{Z}$ puisque $p$ et $q$ sont premiers entre eux. Donc $a\mathbb{Z} + b\mathbb{Z} = \frac{b}{q}\mathbb{Z}$. Cela prouve que $a\mathbb{Z} + b\mathbb{Z}$ est non dense. On a donc démontré que $a\mathbb{Z} + b\mathbb{Z}$ est non dense si et seulement si $\frac{a}{b} \in \mathbb{Q}$. Par contraposition, $a\mathbb{Z} + b\mathbb{Z}$ est dense si et seulement si $\frac{a}{b} \notin \mathbb{Q}$.

(ii) Il faut montrer que pour tout $x \in \mathbb{R}$ et tout $\varepsilon > 0$, il existe $p \in \mathbb{N}$ et $q \in \mathbb{Z}$ tels que

$$|ap + bq - x| < \varepsilon\,. \tag{1.26}$$

Soit donc $x \in \mathbb{R}$ et $\varepsilon > 0$ arbitraires. Comme $a\mathbb{Z} + b\mathbb{Z}$ est dense dans $\mathbb{R}$, il existe $p_0, q_0 \in \mathbb{Z}$ tels que $|ap_0 + bq_0 - x| < \frac{\varepsilon}{2}$. Si $p_0 \in \mathbb{N}$, on a directement (1.26) avec $p = p_0$ et $q = q_0$. Supposons donc $p_0 \notin \mathbb{N}$. Alors $p_0 \leqslant -1$. On va montrer qu'il existe dans ce cas $n, m \in \mathbb{Z}$ avec $n \geqslant 1$ tels que (1.26) soit vérifiée avec $p = p_0 + n|p_0|$ et

$q = q_0 + m|p_0|$. On a bien alors $p \in \mathbb{N}$. La condition (1.26) avec $p = p_0 + n|p_0|$ et $q = q_0 + m|p_0|$ s'écrit

$$|a(p_0 + n|p_0|) + b(q_0 + m|p_0|) - x| < \varepsilon \tag{1.27}$$

*i.e.*

$$|z + w| < \varepsilon \tag{1.28}$$

avec $z = ap_0 + bq_0 - x$ et $w = an|p_0| + bm|p_0|$. Pour que (1.28) soit réalisée, il suffit que $|z| < \frac{\varepsilon}{2}$ et $|w| < \frac{\varepsilon}{2}$. Par définition de $p_0$ et $q_0$ on a $|z| < \frac{\varepsilon}{2}$. La condition $|w| < \frac{\varepsilon}{2}$ s'écrit

$$|an + bm| < \frac{\varepsilon}{2|p_0|}. \tag{1.29}$$

Or puisque $a\mathbb{Z} + b\mathbb{Z}$ est dense dans $\mathbb{R}$, il existe $n, m \in \mathbb{Z}$, non simultanément nuls, tels que $|an + bm| < \frac{\varepsilon}{2|p_0|}$. Changeant éventuellement $n, m$ en leurs opposés, on peut supposer $n \geqslant 0$. Comme $\varepsilon > 0$ est arbitraire, on peut supposer $\varepsilon < b$. Alors on a bien $n \geqslant 1$. En effet, $n = 0$ impliquerait $m \neq 0$ et $b|m| < \frac{\varepsilon}{2|p_0|} < \varepsilon$, donc $b < \frac{\varepsilon}{|m|} \leqslant \varepsilon$. On a donc trouvé $n, m \in \mathbb{Z}$ avec $n \geqslant 1$ tels que (1.29) soit vérifiée, d'où le résultat.

**Exercice 1.11.**

(i) Soit $a = \operatorname{Log} 2$ et $b = \operatorname{Log} 10$. Montrer que $a\mathbb{N} + b\mathbb{Z}$ est dense dans $\mathbb{R}$.

(ii) En déduire qu'il existe une infinité d'entiers $n \in \mathbb{N}$ tels que le développement décimal de $2^n$ commence par un "7". (Le résultat est identique avec n'importe quel chiffre compris entre 1 et 9 au lieu de 7).

**Indication.**

(i) D'après l'exercice 1.10, il suffit de montrer que $\frac{\operatorname{Log} 2}{\operatorname{Log} 10} \notin \mathbb{Q}$. Supposons donc $\frac{\operatorname{Log} 2}{\operatorname{Log} 10} \in \mathbb{Q}$. Il existe alors $p, q \in \mathbb{N}^*$ premiers entre eux tels que $\frac{\operatorname{Log} 2}{\operatorname{Log} 10} = \frac{p}{q}$, donc $q\operatorname{Log} 2 = p\operatorname{Log} 10$ *i.e.* $2^q = 10^p$. Ceci est impossible puisque 5 ne divise pas 2. Donc $\frac{\operatorname{Log} 2}{\operatorname{Log} 10} \notin \mathbb{Q}$.

(ii) Soit $n \in \mathbb{N}$. La première décimale de $2^n$ est un 7 si et seulement si il existe $p \in \mathbb{N}$ tel que

$$7 \cdot 10^p \leqslant 2^n < 8 \cdot 10^p. \tag{1.30}$$

Ceci est équivalent à dire qu'il existe $p \in \mathbb{Z}$ vérifiant (1.30) (puisque $p$ sera alors nécessairement positif) ou encore qu'il existe $p \in \mathbb{Z}$ tel que

$$\operatorname{Log} 7 \leqslant n\operatorname{Log} 2 - p\operatorname{Log} 10 < \operatorname{Log} 8. \tag{1.31}$$

Or d'après (i), $(\operatorname{Log} 2)\,\mathbb{N} + (\operatorname{Log} 10)\,\mathbb{Z}$ est dense dans $\mathbb{R}$, donc il existe une infinité de réels de la forme $n \log 2 - p\operatorname{Log} 10, n \in \mathbb{N}, p \in \mathbb{Z}$, vérifiant (1.31). L'ensemble des $n \in \mathbb{N}$ pour lesquels il existe $p \in \mathbb{Z}$ vérifiant (1.31) est alors aussi infini. En effet, pour chaque $n \in \mathbb{N}$ il existe au plus un tel $p \in \mathbb{Z}$ puisque $\operatorname{Log} 10 > 1$ alors que $\operatorname{Log} 8 - \operatorname{Log} 7 = \operatorname{Log} \frac{8}{7} < 1$.

**Exercice 1.12.**

Montrer que $\{e^{in}; n \in \mathbb{N}\}$ est dense dans $\mathbb{T} = \{z \in \mathbb{C}; |z| = 1\}$.

**Indication.**

$\{e^{in}; n \in \mathbb{N}\}$ est dense dans $\mathbb{T}$ si et seulement si tout $z \in \mathbb{T}$ est limite d'une suite de points de $\{e^{in}; n \in \mathbb{N}\}$. Soit donc $z \in \mathbb{T}$. Il existe $x \in \mathbb{R}$ tel que $z = e^{ix}$. Or

d'après l'exercice 1.10, le sous-groupe $\mathbb{Z} + 2\pi\mathbb{Z}$ est dense dans $\mathbb{R}$ puisque $\pi \notin \mathbb{Q}$, et cela implique que le sous-ensemble $\mathbb{N}+2\pi\mathbb{Z}$ aussi est dense dans $\mathbb{R}$. Il existe donc une suite de points $x_k = n_k + 2p_k\pi \in \mathbb{N}+2\pi\mathbb{Z}$ telle que $x = \lim_k x_k$. L'application $t \mapsto e^{it}$ de $\mathbb{R}$ dans $\mathbb{T}$ étant continue, on a $z = e^{ix} = \lim_k e^{ix_k}$. Or $e^{ix_k} = e^{in_k}e^{2i\pi p_k} = e^{in_k}$. Donc $z = \lim_k e^{in_k}$.

**Exercice 1.13.**

Soit $G$ un groupe et $x$ un élément d'ordre $n > 1$ de $G$. Quel est l'ordre de $x^p$, $p \in \mathbb{N}, p > 1$ ?

**Indication.**

L'ordre de $x^p$ est le plus petit entier strictement positif $k \in \mathbb{N}^*$ tel que $x^{pk} = e$. Or $x^{pk} = e$ si et seulement si $pk$ est divisible par $n$. Soit $D = p \curlywedge n$. On a $n = Dn_1$, $p = Dp_1$ avec $n_1 \curlywedge p_1 = 1$. Alors :

$$\begin{aligned} pk \text{ divisible par } n &\Leftrightarrow \exists q \in \mathbb{N} \quad pk = nq \\ &\Leftrightarrow \exists q \in \mathbb{N} \quad p_1 k = n_1 q \\ &\Leftrightarrow k \text{ divisible par } n_1 \text{ (Th. de Gauss)} . \end{aligned}$$

Donc l'ordre de $x^p$ est $n_1 = \frac{n}{D}$.

**Exercice 1.14.**

Soit $G$ un groupe abélien, $e$ son élément neutre, $x$ et $y$ des éléments de $G$ d'ordre respectifs finis $p$ et $q$.

(i) Montrer que si $\langle x \rangle \cap \langle y \rangle = \{e\}$, alors $xy$ est d'ordre $p \curlyvee q$, où $p \curlyvee q$ dénote le PPCM de $p$ et $q$.

(ii) En déduire que si $p \curlywedge q = 1$, alors $xy$ est d'ordre $p \curlyvee q$.

(iii) Dans le groupe $\mathbb{Z}_4 \times \mathbb{Z}_6$, quels sont les ordres de $x = ([1],[0])$ et $y = ([1],[2])$? Vérifier que l'ordre de $x+y$ n'est pas le PPCM des ordres de $x$ et $y$. Vérifier que $\langle x \rangle \cap \langle y \rangle \neq \{([0],[0])\}$.

(iv) Dans le groupe $\mathcal{S}_3$ trouver un élément $\sigma$ d'ordre 3 et un élément $\tau$ d'ordre 2 tels que $\sigma\tau$ soit d'ordre 2.

**Indication.**

(i) L'ordre de $xy$ est le plus petit entier strictement positif $k \in \mathbb{N}^*$ tel que $(xy)^k = e$. Or $(xy)^k = x^k y^k$ puisque $G$ est abélien, et l'on a :

$$\begin{aligned} x^k y^k = e &\Leftrightarrow x^k = y^{-k} \\ &\Leftrightarrow x^k = e \text{ et } y^k = e \text{ (puisque } x^k = y^{-k} \in \langle x \rangle \cap \langle y \rangle = \{e\} \text{ )} \\ &\Leftrightarrow k \in p\mathbb{Z} \text{ et } k \in q\mathbb{Z} \\ &\Leftrightarrow k \in p\mathbb{Z} \cap q\mathbb{Z} = (p \curlyvee q)\mathbb{Z} \end{aligned}$$

Donc l'ordre de $xy$ est $p \curlyvee q$.

(ii) $\langle x \rangle \cap \langle y \rangle$ est un sous-groupe de $\langle x \rangle$ donc son cardinal divise le cardinal de $\langle x \rangle$ qui est $p$. De même, il divise le cardinal de $\langle y \rangle$ qui est $q$. Comme $p \curlywedge q = 1$, le cardinal de $\langle x \rangle \cap \langle y \rangle$ est donc 1, *i.e.* $\langle x \rangle \cap \langle y \rangle = \{e\}$. D'où le résultat d'après la question (i).

(iii) On a $y = ([1],[0]) + ([0],[2])$. Or $x = ([1],[0])$ est d'ordre 4, et $([0],[2])$ est d'ordre 3. Comme 4 et 3 sont premiers entre eux, $y$ est d'ordre $4 \curlyvee 3 = 12$. Considérons $x + y = ([2],[2])$. On a aussi $x + y = ([2],[0]) + ([0],[2])$. Or $([2],[0])$ est d'ordre 2 et $([0],[2])$ est d'ordre 3. Comme 2 est premier avec 3, $x+y$ est donc d'ordre 6. On constate que ce n'est pas le PPCM des ordres de $x$ et de $y$ : $6 \neq 4 \curlyvee 12 = 12$. On a en explicitant tous les éléments de $\langle y \rangle$ :

$$\langle x \rangle \cap \langle y \rangle = \{([0],[0]), ([1],[0]), ([2],[0]), ([3],[0])\} = \mathbb{Z}_4 \times \{[0]\}.$$

(iv) On peut prendre $\sigma = (1,2,3)$ et $\tau = (1,2)$. On a $\sigma\tau = (1,3)$. Le point ici est que le groupe $\mathcal{S}_3$ n'est pas commutatif.

**Exercice 1.15.**

(i) Soit $G$ un groupe abélien fini et $m$ le PPCM des ordres de *tous* les éléments de $G$. Montrer qu'il existe un élément $x \in G$ d'ordre $m$.

(ii) En déduire que pour $p$ nombre premier quelconque, le groupe multiplicatif $(\mathbb{Z}\,/\,p\mathbb{Z})^*$ des éléments non nuls du corps $\mathbb{Z}\,/\,p\mathbb{Z}$ est un groupe cyclique isomorphe à $\mathbb{Z}_{p-1}$.

**Indication.**

(i) Soient $x_1, \ldots, x_n$ les divers éléments de $G$, et $\omega_1, \ldots, \omega_n$ leurs ordres. Considérons la décomposition en facteurs premiers du PPCM $m$ :

$$m = p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_r^{\alpha_r} = \prod_{\lambda=1}^{r} p_\lambda^{\alpha_\lambda}.$$

Fixons un indice $\lambda$. Pour chaque $\omega_i$ $(1 \leqslant i \leqslant n)$ soit $k_i$ la puissance la plus élevée telle que $p_\lambda^{k_i}$ divise $\omega_i$. Alors par définition du PPCM, $\alpha_\lambda = \sup(k_1, \ldots, k_n)$. Il existe donc un indice $i_\lambda$ tel que $k_{i_\lambda} = \alpha_\lambda$ , *i.e.* $\omega_{i_\lambda} = q_\lambda p_\lambda^{\alpha_\lambda}$, $q_\lambda \curlywedge p_\lambda = 1$. D'après l'exercice 1.13, $x_{i_\lambda}^{q_\lambda}$ a pour ordre $\frac{\omega_{i_\lambda}}{\omega_{i_\lambda} \curlywedge q_\lambda} = \frac{q_\lambda p_\lambda^{\alpha_\lambda}}{q_\lambda} = p_\lambda^{\alpha_\lambda}$. Maintenant, d'après l'exercice 1.14, $x_{i_1}^{q_1} x_{i_2}^{q_2}$ a pour ordre $p_1^{\alpha_1} p_2^{\alpha_2}$ et par récurrence, $x = x_{i_1}^{q_1} x_{i_2}^{q_2} \cdots x_{i_r}^{q_r}$ a pour ordre $p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r} = m$.

(ii) $(\mathbb{Z}\,/\,p\mathbb{Z})^*$ est un groupe abélien d'ordre $p-1$. Supposons qu'il n'existe pas d'élément d'ordre $p-1$. Alors d'après la question précédente, le PPCM $m$ des ordres des éléments de $(\mathbb{Z}\,/\,p\mathbb{Z})^*$ est strictement plus petit que $p-1$. Or on a $x^m = 1\ \forall x \in (\mathbb{Z}\,/\,p\mathbb{Z})^*$. Le polynôme $X^m - 1 \in (\mathbb{Z}\,/\,p\mathbb{Z})[X]$ a donc au moins pour racines dans le corps *commutatif* $\mathbb{Z}\,/\,p\mathbb{Z}$ les $p-1$ éléments de $(\mathbb{Z}\,/\,p\mathbb{Z})^*$. Mais il est de degré $m$ et a donc au plus $m$ racines dans le corps $\mathbb{Z}\,/\,p\mathbb{Z}$. Cela implique $m \geqslant p-1$, ce qui est contradictoire. L'hypothèse de départ est donc fausse, *i.e.* il existe dans $(\mathbb{Z}\,/\,p\mathbb{Z})^*$ un élément d'ordre $p-1$, d'où le résultat.

**Exercice 1.16.**

Trouver tous les homomorphismes de corps de $\mathbb{R}$ dans lui-même.

**Indication.**

Soit $f$ un homomorphisme de corps de $\mathbb{R}$ dans lui-même. On a $f(1) = 1$, donc $f(n) = f(\underbrace{1 + 1 + \cdots + 1}_{n \text{ fois}}) = \underbrace{f(1) + f(1) + \cdots + f(1)}_{n \text{ fois}}) = n\,\forall n \in \mathbb{N}^*$. Mais $f(-n) = -f(n)$ donc $f(n) = n\,\forall n \in \mathbb{Z}$. Ensuite $1 = f(1) = f(n\frac{1}{n}) = f(n)f(\frac{1}{n})$ donc $f(\frac{1}{n}) = \frac{1}{n}\,\forall n \in \mathbb{N}^*$ et enfin $f(\frac{p}{q}) = f(p)f(\frac{1}{q}) = \frac{p}{q}\,\forall p \in \mathbb{Z}, \forall q \in \mathbb{N}^*$, *i.e.* $f(r) = r\ \forall r \in \mathbb{Q}$. D'autre part $f(x) = f(\sqrt{x})f(\sqrt{x}) \geqslant 0\ \forall x \geqslant 0$. Si $x, y \in \mathbb{R}$ sont tels que $x \leqslant y$, on a $y - x \geqslant 0$, donc $f(y-x) \geqslant 0$, ce qui donne $f(x) \leqslant f(y)$. Ainsi l'application $f$ est croissante. Si $x \in \mathbb{R} \setminus \mathbb{Q}$, comme $\mathbb{Q}$ est dense dans $\mathbb{R}$, il existe 2 suites de rationnels $(\alpha_k)$ et $(\beta_k)$ telles que $\alpha_k < x < \beta_k\,\forall k$ et $x = \lim_k \alpha_k = \lim_k \beta_k$. Cela implique puisque $f$ est croissante $f(\alpha_k) = \alpha_k \leqslant f(x) \leqslant f(\beta_k) = \beta_k\,\forall k$ d'où $f(x) = x$ par passage à la limite quand $k \to +\infty$. On a donc $f = Id_{\mathbb{R}}$. Le seul homorphisme de corps de $\mathbb{R}$ dans lui-même est l'identité.

**Exercice 1.17.**

Soit $\mathcal{A}$ un anneau commutatif unitaire *fini*. Montrer qu'un élément $a \in \mathcal{A}$ est inversible pour la multiplication si et seulement si ce n'est pas un diviseur de 0, *i.e.* $ab \neq 0\ \forall b \neq 0$.

**Indication.**

Si $a$ est inversible, $ab = 0$ implique $b = a^{-1} \cdot 0 = 0$, donc $a$ n'est pas un diviseur de 0. Réciproquement, soit $a \in \mathcal{A}$ et supposons que $a$ ne soit pas un diviseur de 0. Considérons l'application $f : \mathcal{A} \to \mathcal{A}$ définie par $f(x) = ax \ \forall x \in \mathcal{A}$. Si $x, x' \in \mathcal{A}$, on a $f(x) = f(x') \Leftrightarrow a(x - x') = 0 \Leftrightarrow x - x' = 0$ car $a$ n'est pas un diviseur de 0. Donc $f$ est injective. Mais $\mathcal{A}$ étant fini, $f$ est alors bijective. Il existe par suite un $x \in \mathcal{A}$ tel que $f(x) = 1$. Ainsi $a$ est inversible et son inverse est $x$.

**Exercice 1.18.**

Soit $\mathcal{A}$ un anneau commutatif unitaire et $I$ un idéal de $\mathcal{A}$. On pose:

$$\sqrt{I} = \{x \in \mathcal{A};\ \exists n \in \mathbb{N} \quad x^n \in I\}.$$

(i) Montrer que $\sqrt{I}$ est un idéal contenant $I$.
(ii) Si $I, J$ sont deux idéaux tels que $I \subset J$, montrer que $\sqrt{I} \subset \sqrt{J}$.
(iii) Montrer que $\sqrt{\sqrt{I}} = \sqrt{I}$.
(iv) Montrer que si $I, J$ sont deux idéaux on a $\sqrt{I} \cap \sqrt{J} = \sqrt{I \cap J}$.

**Exercice 1.19.**

Soit $X$ un ensemble fini, $\mathcal{A}$ l'ensemble des fonctions à valeurs complexes sur $X$.
(i) Montrer que $\mathcal{A}$ est un anneau commutatif.
(ii) Trouver les idéaux de $\mathcal{A}$. En déduire que $\mathcal{A}$ est principal.
(iii) Trouver les idéaux maximaux de $\mathcal{A}$.

**Indication.**

(ii) Pour $Z \subset X$, introduisons $\mathcal{I}_Z = \{f \in \mathcal{A};\ f(x) = 0 \quad \forall x \in Z\}$. Il est immédiat que $\mathcal{I}_Z$ est un idéal de $\mathcal{A}$. On va montrer que tout idéal de $\mathcal{A}$ est de ce type, *i.e.* pour tout idéal $\mathcal{J}$ de $\mathcal{A}$, il existe un sous-ensemble $Z \subset X$ tel que $\mathcal{J} = \mathcal{I}_Z$. Soit donc $\mathcal{J}$ un idéal de $\mathcal{A}$. Soit $Z = \{x \in X;\ f(x) = 0 \quad \forall f \in \mathcal{J}\}$ l'ensemble des points de $X$ où s'annulent toutes les fonctions appartenant à $\mathcal{J}$. On a alors par définition de $\mathcal{I}_Z$ :

$$\mathcal{J} \subset \mathcal{I}_Z. \tag{1.32}$$

Il reste à montrer que $\mathcal{I}_Z \subset \mathcal{J}$. Pour cela, on va montrer que la fonction $F = \mathbf{1}_{X \setminus Z}$ définie par

$$F(x) = \begin{cases} 1 & \text{si } x \notin Z \\ 0 & \text{si } x \in Z \end{cases}$$

appartient à $\mathcal{J}$. On aura alors pour tout $g \in \mathcal{I}_Z$

$$g = gF \in \mathcal{A}F = (F) \subset \mathcal{J}$$

donc

$$\mathcal{I}_Z \subset (F) \subset \mathcal{J}.$$

D'après (1.32), on aura ainsi

$$\mathcal{I}_Z = (F) = \mathcal{J}.$$

Cela montre en particulier que $\mathcal{A}$ est un anneau principal.

Montrons donc maintenant que la fonction $F$ appartient à $\mathcal{J}$. Si $Z = X$, la seule fonction de $\mathcal{J}$ est la fonction nulle, donc $\mathcal{J} = \{0\}$. Or si $Z = X$, la fonction $F$ est nulle, donc on a bien $F \in \mathcal{J}$. On peut donc supposer $Z \neq X$. Alors pour tout $y \notin Z$

il existe par définition de $Z$ une fonction $f_y \in \mathcal{J}$ telle que $f_y(y) \neq 0$. Considérons la fonction $\mathbf{1}_{\{y\}}$ définie par

$$\mathbf{1}_{\{y\}}(x) = \begin{cases} 1 & \text{si } x = y \\ 0 & \text{si } x \neq y \end{cases}$$

et soit $h_y \in \mathcal{A}$ définie par

$$h_y = \frac{1}{f_y(y)} \mathbf{1}_{\{y\}}.$$

Alors $h_y f_y = \mathbf{1}_{\{y\}}$. Comme $f_y \in \mathcal{J}$, cela implique $\mathbf{1}_{\{y\}} \in \mathcal{J}$. Or $F = \sum_{y \notin Z} \mathbf{1}_{\{y\}}$. Donc $F \in \mathcal{J}$.
(iii) Pour $Z_1, Z_2 \subset X$, on voit facilement que

$$\mathcal{I}_{Z_1} \subset \mathcal{I}_{Z_2} \iff Z_1 \supset Z_2. \tag{1.33}$$

Comme $\mathcal{I}_\varnothing = \mathcal{A}$, (1.33) montre qu'un idéal $\mathcal{J} = \mathcal{I}_Z$ de $\mathcal{A}$ est maximal si et seulement si $Z$ est réduit à un élément : $\exists a \in X \quad Z = \{a\}$.

**Exercice 1.20.**

Soit $\mathbb{H} = \left\{ \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \in M_2(\mathbb{C})\,;\ \alpha, \beta \in \mathbb{C} \right\}$.

(i) Montrer que $\mathbb{H}$ est un corps. Les éléments de $\mathbb{H}$ s'appellent des *quaternions*. On notera $e_0, e_1, e_2, e_3$ les quaternions suivants :

$$e_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},\ e_1 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix},\ e_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},\ e_3 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

(ii) Montrer que l'application $f : \mathbb{C} \to \mathbb{H}$ définie par $\lambda \mapsto f(\lambda) = \begin{pmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{pmatrix}$ est un isomorphisme du corps $\mathbb{C}$ sur un sous-corps de $\mathbb{H}$.

(iii) Vérifier que dans $\mathbb{H}$

$$f(\lambda) e_2 = e_2 f(\bar{\lambda}) \quad \forall \lambda \in \mathbb{C}. \tag{1.34}$$

En déduire que le corps $\mathbb{H}$ n'est pas commutatif.

(iv) Soit $Q = \{\pm e_0, \pm e_1, \pm e_2, \pm e_3\} \subset \mathbb{H}$. Montrer que $Q$ est un groupe pour la multiplication des quaternions. $Q$ est-il commutatif?

(v) Dans le corps non commutatif $\mathbb{H}$, citer 6 solutions différentes de l'équation $x^2 = -e_0$ et 8 solutions différentes de l'équation $x^4 = e_0$.

(vi) Montrer que $\mathbb{H}$ est un $\mathbb{C}$-espace vectoriel et que $(e_0, e_2)$ en est une base. Montrer que $(e_0, e_1, e_2, e_3)$ est une base du $\mathbb{R}$-espace vectoriel $\mathbb{H}$.

(vii) Pour $q = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \in \mathbb{H}$, $(\alpha, \beta \in \mathbb{C})$, on pose $\|q\|^2 = \det q$. Comment faut-il définir le conjugué quaternionique $q^*$ de $q$ en sorte que l'on ait la formule $q^{-1} = f(\frac{1}{\|q\|^2}) q^*$ pour $q \neq 0$? Vérifier qu'on a alors $f(\|q\|^2) = qq^*$, $(qq')^* = (q')^* q^*$ pour tous $q, q' \in \mathbb{H}$, et $(f(\lambda))^* = f(\bar{\lambda})$ pour $\lambda \in \mathbb{C}$. Calculer $\|q\|^2$ et $q^*$ pour $q = x^0 \cdot e_0 + x^1 \cdot e_1 + x^2 \cdot e_2 + x^3 \cdot e_3$, $(x^0, x^1, x^2, x^3 \in \mathbb{R})$ dans le $\mathbb{R}$-espace vectoriel $\mathbb{H}$ défini en (vi).

(viii) Montrer que le centre $Z(\mathbb{H})$ du corps $\mathbb{H}$ est $\mathbb{R} \cdot e_0 = f(\mathbb{R})$.

**Indication.**

(i) $\mathbb{H}$ est un sous-anneau de $M_2(\mathbb{C})$. On vérifie en effet facilement que la matrice nulle 0 appartient à $\mathbb{H}$, et que pour tous $q, q' \in \mathbb{H}$ on a $q + q', -q, qq' \in \mathbb{H}$. La matrice identité $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ est l'élément $e_0$ de $\mathbb{H}$, donc $e_0 = 1_{\mathbb{H}}$. Pour vérifier que $\mathbb{H}$ est un corps, il reste à montrer que tout $q \in \mathbb{H}$, $q \neq 0$ est inversible dans $\mathbb{H}$. Soit donc $q = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \neq 0$. Alors $\alpha$ et $\beta$ ne sont pas simultanément nuls, donc $\det q = |\alpha|^2 + |\beta|^2 \neq 0$. La matrice $q$ est inversible. La comatrice de la matrice $q$ étant $\begin{pmatrix} \bar{\alpha} & \bar{\beta} \\ -\beta & \alpha \end{pmatrix}$, on a dans $M_2(\mathbb{C})$

$$q^{-1} = \frac{1}{\det q} {}^t\begin{pmatrix} \bar{\alpha} & \bar{\beta} \\ -\beta & \alpha \end{pmatrix} = \begin{pmatrix} \frac{1}{|\alpha|^2+|\beta|^2}\bar{\alpha} & -\frac{1}{|\alpha|^2+|\beta|^2}\beta \\ \frac{1}{|\alpha|^2+|\beta|^2}\bar{\beta} & \frac{1}{|\alpha|^2+|\beta|^2}\alpha \end{pmatrix}.$$

On voit que $q^{-1} \in \mathbb{H}$.

(ii) On a $f(\lambda + \mu) = f(\lambda) + f(\mu)$, $f(\lambda\mu) = f(\lambda)f(\mu)$, $f(1) = e_0$. Donc $f$ est un homorphisme du corps $\mathbb{C}$ dans le corps $\mathbb{H}$. D'après le Th. 1.19, $f$ est un isomorphisme du corps $\mathbb{C}$ sur le sous-corps Im $f$ de $\mathbb{H}$. On notera que $f(1) = e_0$ et $f(i) = e_1$. Si l'on identifie $\mathbb{C}$ au sous-corps Im $f$, 1 est identifié à $e_0$ et $i$ à $e_1$.

(iii) L'équation (1.34) s'écrit

$$\begin{pmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} \bar{\lambda} & 0 \\ 0 & \lambda \end{pmatrix}.$$

Sa vérification est immédiate. Prenant $\lambda = i$, on obtient $e_1e_2 = -e_2e_1$ donc $\mathbb{H}$ n'est pas commutatif.

(iv) Il suffit de montrer que $Q$ est un sous-groupe du groupe multiplicatif $\mathbb{H}^*$ des éléments non nuls du corps $\mathbb{H}$. On a $Q \subset \mathbb{H}^*$, et l'élément neutre $e_0$ de la multiplication appartient à $Q$. Pour vérifier que $Q$ est stable pour la multiplication, on calcule les divers produits d'abord sans tenir compte du signe.

$e_1^2 = -e_0$.

$e_1e_2 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} = e_3.$

$e_1e_3 = e_1e_1e_2 = e_1^2e_2 = -e_2$

$e_2e_1 = -e_1e_2$ (d'après (1.34)) $= -e_3$

$e_2^2 = -e_0$.

$e_2e_3 = e_2e_1e_2 = -e_1e_2e_2$ (d'après (1.34)) $= e_1$

$e_3e_1 = e_1e_2e_1 = -e_1e_1e_2$ (d'après (1.34)) $= e_2$

$e_3e_2 = e_1e_2e_2 = -e_1$

$e_3^2 = e_1e_2e_3 = e_1^2 = -e_0$.

Cela donne le tableau suivant :

| $\times$ | $e_0$ | $e_1$ | $e_2$ | $e_3$ |
|---|---|---|---|---|
| $e_0$ | $e_0$ | $e_1$ | $e_2$ | $e_3$ |
| $e_1$ | $e_1$ | $-e_0$ | $e_3$ | $-e_2$ |
| $e_2$ | $e_2$ | $-e_3$ | $-e_0$ | $e_1$ |
| $e_3$ | $e_3$ | $e_2$ | $-e_1$ | $-e_0$ |

On en déduit immédiatement, en prenant maintenant en compte aussi les signes, que $Q$ est stable par multiplication. On en déduit aussi que $Q$ est stable par passage

à l'inverse : l'inverse de $\pm e_0$ est $\pm e_0$, celui de $\pm e_k$ est $\mp e_k$ pour $k = 1, 2, 3$.
(v) $\pm e_1, \pm e_2, \pm e_3$ sont 6 solutions différentes de l'équation $x^2 = -e_0$. Les 8 éléments de $Q$ sont solutions de l'équation $x^4 = e_0$.
(vi) Notons d'abord que $\mathbb{H}$ *n'est pas* un sous-espace vectoriel de $M_2(\mathbb{C})$ pour sa structure habituelle.

En effet, si $q = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \in \mathbb{H}$, $q \neq 0$ $(\alpha, \beta \in \mathbb{C})$, on a $\begin{pmatrix} i\alpha & i\beta \\ -i\bar{\beta} & i\bar{\alpha} \end{pmatrix} \notin \mathbb{H}$.

Maintenant, $\mathbb{H}$ est un corps contenant un sous-corps isomorphe à $\mathbb{C}$. Il est donc immédiat que c'est un $\mathbb{C}$-espace vectoriel pour l'addition $(q, q') \in \mathbb{H} \times \mathbb{H} \mapsto q + q' \in \mathbb{H}$ et la multiplication dans $\mathbb{H}$ par un complexe $\lambda$ : $(\lambda, q) \in \mathbb{C} \times \mathbb{H} \mapsto \lambda \cdot q = f(\lambda)q \in \mathbb{H}$. On prendra garde que $\lambda \cdot q$ *n'est pas* le produit usuel du complexe $\lambda$ par la matrice $q$ dans $M_2(\mathbb{C})$. C'est le produit usuel seulement si $\lambda \in \mathbb{R}$ puisqu'alors $f(\lambda) = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$.
C'est pour cette raison de contexte matriciel que nous n'effectuons pas dans ce qui suit l'identification entre $\mathbb{C}$ et $\mathrm{Im}\, f$. Notons aussi que
$(\lambda \cdot q)q' = (f(\lambda)q)q' = f(\lambda)(qq') = \lambda \cdot (qq')$ pour $\lambda \in \mathbb{C}$, $q, q' \in \mathbb{H}$.

Tout $q = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \in \mathbb{H}$, $(\alpha, \beta \in \mathbb{C})$ s'écrit

$$q = \begin{pmatrix} \alpha & 0 \\ 0 & \bar{\alpha} \end{pmatrix} + \begin{pmatrix} 0 & \beta \\ -\bar{\beta} & 0 \end{pmatrix} = \begin{pmatrix} \alpha & 0 \\ 0 & \bar{\alpha} \end{pmatrix} e_0 + \begin{pmatrix} \beta & 0 \\ 0 & \bar{\beta} \end{pmatrix} e_2 = f(\alpha)e_0 + f(\beta)e_2,$$

*i.e.* $q = \alpha \cdot e_0 + \beta \cdot e_2$. De plus, si $q = \alpha' \cdot e_0 + \beta' \cdot e_2$, on a nécessairement $\alpha = \alpha'$ et $\beta = \beta'$, *i.e.* la décomposition est unique. Donc $(e_0, e_2)$ est une base du $\mathbb{C}$-espace vectoriel $\mathbb{H}$. De même $q$ se décompose de façon unique sous la forme

$$q = \begin{pmatrix} x^0 & 0 \\ 0 & x^0 \end{pmatrix} + \begin{pmatrix} ix^1 & 0 \\ 0 & -ix^1 \end{pmatrix} + \begin{pmatrix} 0 & x^2 \\ -x^2 & 0 \end{pmatrix} + \begin{pmatrix} 0 & ix^3 \\ ix^3 & 0 \end{pmatrix},$$

*i.e.*

$$q = x^0 \cdot e_0 + x^1 \cdot e_1 + x^2 \cdot e_2 + x^3 \cdot e_3$$

avec $x^0 = \mathfrak{Re}\,(\alpha)$, $x^1 = \mathfrak{Im}\,(\alpha)$, $x^2 = \mathfrak{Re}\,(\beta)$, $x^3 = \mathfrak{Im}\,(\beta)$, donc $(e_0, e_1, e_2, e_3)$ est une base du $\mathbb{R}$-espace vectoriel $\mathbb{H}$.
(vii) On a vu que pour $q \neq 0$

$$q^{-1} = \begin{pmatrix} \frac{1}{|\alpha|^2+|\beta|^2}\bar{\alpha} & -\frac{1}{|\alpha|^2+|\beta|^2}\beta \\ \frac{1}{|\alpha|^2+|\beta|^2}\bar{\beta} & \frac{1}{|\alpha|^2+|\beta|^2}\alpha \end{pmatrix}.$$

Cela s'écrit dans $\mathbb{H}$

$$q^{-1} = f(\frac{1}{\|q\|^2})q^* = \frac{1}{\|q\|^2} \cdot q^*$$

si l'on pose $q^* = \begin{pmatrix} \bar{\alpha} & -\beta \\ \bar{\beta} & \alpha \end{pmatrix}$ pour tout $q = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \in \mathbb{H}$ . A noter que la matrice $q^*$ n'est autre que la transconjuguée (*i.e.* la transposée conjuguée $\overline{{}^t q} = {}^t\bar{q}$) de la matrice $q$. On a

$$qq^* = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \begin{pmatrix} \bar{\alpha} & -\beta \\ \bar{\beta} & \alpha \end{pmatrix} = \begin{pmatrix} \det q & 0 \\ 0 & \det q \end{pmatrix} = f(\|q\|^2) = \|q\|^2 \cdot e_0 .$$

La relation $(qq')^* = (q')^* q^* \quad \forall\, q, q' \in \mathbb{H}$ se vérifie facilement :

$$(qq')^* = \overline{{}^t(qq')} = \overline{{}^tq'\,{}^tq} = \overline{{}^tq'}\;\overline{{}^tq} = (q')^* q^*.$$

Pour $\lambda \in \mathbb{C}$ on a $f(\lambda)^* = \begin{pmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{pmatrix}^* = \begin{pmatrix} \bar{\lambda} & 0 \\ 0 & \lambda \end{pmatrix} = f(\bar{\lambda})$. Enfin

$$q^* = (\alpha \cdot e_0 + \beta \cdot e_2)^* = \bar{\alpha} \cdot e_0 - \beta \cdot e_2 = x^0 \cdot e_0 - x^1 \cdot e_1 - x^2 \cdot e_2 - x^3 \cdot e_3 .$$

$$\|q\|^2 = |\alpha|^2 + |\beta|^2 = (x^0)^2 + (x^1)^2 + (x^2)^2 + (x^3)^2 .$$

(viii) D'abord il est clair que $f(\mathbb{R}) = \mathbb{R} \cdot e_0 \subset Z(\mathbb{H})$. En effet, si $\lambda \in \mathbb{R}$, $\lambda \cdot e_0 = f(\lambda)$ est la matrice $\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$ qui commute avec toute matrice. Maintenant, soit $q = x^0 \cdot e_0 + x^1 \cdot e_1 + x^2 \cdot e_2 + x^3 \cdot e_3 \in \mathbb{H}$ $(x^0, x^1, x^2, x^3 \in \mathbb{R})$, et supposons $q \in Z(\mathbb{H})$. On a

$$qe_1 = x^0 \cdot e_1 - x^1 \cdot e_0 - x^2 \cdot e_3 + x^3 \cdot e_2,$$

$$e_1 q = x^0 \cdot e_1 - x^1 \cdot e_0 + x^2 \cdot e_3 - x^3 \cdot e_2.$$

Or $qe_1 = e_1 q$ donc $x_2 = x_3 = 0$. De même, avec $e_2$,

$$qe_2 = (x^0 \cdot e_0 + x^1 \cdot e_1)e_2 = x^0 \cdot e_2 + x^1 \cdot e_3,$$

$$e_2 q = e_2(x^0 \cdot e_0 + x^1 \cdot e_1) = x^0 \cdot e_2 - x^1 \cdot e_3.$$

Or $qe_2 = e_2 q$ donc $x_1 = 0$ et $q = x^0 \cdot e_0 \in \mathbb{R} \cdot e_0$. Cela prouve que $Z(\mathbb{H}) \subset \mathbb{R} \cdot e_0$, d'où $Z(\mathbb{H}) = \mathbb{R} \cdot e_0$.

**Exercice 1.21.**

(i) Trouver une matrice $V \in M_2(\mathbb{C})$ telle que l'on ait *pour tous* $\alpha, \beta \in \mathbb{C}$ :

$$V \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} V^{-1} = \begin{pmatrix} \bar{\alpha} & \bar{\beta} \\ -\beta & \alpha \end{pmatrix}.$$

(ii) Montrer qu'il n'existe pas de matrice $W \in M_2(\mathbb{C})$ telle que l'on ait *pour tous* $\alpha, \beta \in \mathbb{C}$ :

$$W \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} W^{-1} \in M_2(\mathbb{R}).$$

**Indication.**

(i) $V = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ convient.

(ii) La matrice $\begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}$ $(\alpha, \beta \in \mathbb{C})$ est un quaternion $q$ (voir ex. 1.20). Supposons qu'il existe une matrice $W \in M_2(\mathbb{C})$ telle que

$$W \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} W^{-1} \in M_2(\mathbb{R}) \quad \forall \alpha, \beta \in \mathbb{C}. \tag{1.35}$$

On a en particulier en prenant $q = e_1 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$ :

$$\widetilde{e_1} = W e_1 W^{-1} \in M_2(\mathbb{R}).$$

Or $\operatorname{Tr} \widetilde{e_1} = \operatorname{Tr} e_1 = 0$. On peut donc écrire

$$\widetilde{e_1} = \begin{pmatrix} a & b+c \\ b-c & -a \end{pmatrix}$$

avec $a, b, c \in \mathbb{R}$. Maintenant

$$\widetilde{e_1}^2 = (We_1W^{-1})^2 = We_1{}^2W^{-1} = W\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}W^{-1} = \begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}.$$

Or

$$\widetilde{e_1}^2 = \begin{pmatrix}a & b+c\\ b-c & -a\end{pmatrix}^2 = \begin{pmatrix}a^2+b^2-c^2 & 0\\ 0 & a^2+b^2-c^2\end{pmatrix}$$

donc

$$a^2 + b^2 - c^2 = -1. \tag{1.36}$$

De même, en prenant $q = e_2 = \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}$, on obtient :

$$\widetilde{e_2} = We_2W^{-1} = \begin{pmatrix}a' & b'+c'\\ b'-c' & -a'\end{pmatrix}$$

avec $a', b', c' \in \mathbb{R}$ et

$$a'^2 + b'^2 - c'^2 = -1. \tag{1.37}$$

Maintenant,

$$\mathrm{Tr}\,(\widetilde{e_1}\widetilde{e_2}) = \mathrm{Tr}\,(We_1W^{-1}We_2W^{-1}) = \mathrm{Tr}\,(We_1e_2W^{-1}) = \mathrm{Tr}\,(e_1e_2) = \mathrm{Tr}\,e_3 = 0,$$

où $e_3 = \begin{pmatrix}0 & i\\ i & 0\end{pmatrix}$. Or

$$\widetilde{e_1}\widetilde{e_2} = \begin{pmatrix}aa' + (b+c)(b'-c') & a(b'+c') - a'(b+c)\\ a'(b-c) - a(b'-c') & aa' + (b-c)(b'+c')\end{pmatrix}$$

donc $\mathrm{Tr}\,(\widetilde{e_1}\widetilde{e_2}) = 2(aa' + bb' - cc')$ et par conséquent

$$aa' + bb' - cc' = 0. \tag{1.38}$$

D'après l'inégalité de Schwarz, on a alors compte tenu de (1.36) et (1.37)

$$cc' = aa' + bb' \leqslant \sqrt{a^2+b^2}\sqrt{a'^2+b'^2} = \sqrt{c^2-1}\sqrt{c'^2-1}$$

d'où par élévation au carré

$$c^2 + c'^2 \leqslant 1.$$

Cela est en contradiction avec (1.36) et (1.37). Il n'existe donc pas de W vérifiant (1.35).

# Chapitre 2

# Groupe orthogonal, Groupe euclidien.

## 2.1 Rappels.

### 2.1.1 Vecteurs et matrices.

**Bases d'un espace vectoriel.**

Soit $K$ un corps commutatif et $E$ un $K$-espace vectoriel. Un système de vecteurs $\mathcal{B} = (\mathbf{f_1}, \ldots, \mathbf{f_n})$ est une base de $E$ si tout vecteur $\mathbf{x} \in E$ s'écrit de façon unique $\mathbf{x} = \sum_{j=1}^{n} \xi^j \mathbf{f_j} \quad (\xi^j \in K \quad \forall j)$. La matrice colonne $X = \begin{pmatrix} \xi^1 \\ \vdots \\ \xi^n \end{pmatrix}$ est la matrice des composantes du vecteur $\mathbf{x}$ dans la base $\mathcal{B}$. Si $E$ possède une base ayant $n$ éléments, toute autre base de $E$ a aussi $n$ éléments. On dit que $E$ est de dimension finie $n$ et on note $\dim E = n$.

**Base canonique de $K^n$.**

Si $E$ est l'espace vectoriel $K^n$, un vecteur $\mathbf{x} \in E$ est une suite de $n$ éléments de $K$ notée en colonne $\mathbf{x} = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}$. Les vecteurs $\mathbf{e_1} = \begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \mathbf{e_n} = \begin{pmatrix} 0 \\ \vdots \\ 1 \end{pmatrix}$ forment la base canonique $(\mathbf{e_1}, \ldots, \mathbf{e_n})$ de $K^n$. Le vecteur $\mathbf{x} = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}$ s'identifie à sa matrice colonne dans la base canonique.

**Matrice de passage.**

Soit $E$ un $K$-espace vectoriel de dimension finie $n$. Si $\mathcal{B} = (\mathbf{f_1}, \ldots, \mathbf{f_n})$ et $\mathcal{B}' = (\mathbf{f'_1}, \ldots, \mathbf{f'_n})$ sont deux bases de $E$, la matrice $P = \begin{pmatrix} \alpha_1^1 & \ldots & \alpha_n^1 \\ \vdots & \ldots & \vdots \\ \alpha_1^n & \ldots & \alpha_n^n \end{pmatrix}$, où $\begin{pmatrix} \alpha_j^1 \\ \vdots \\ \alpha_j^n \end{pmatrix}$ est la matrice colonne des composantes du vecteur $\mathbf{f'_j}$ dans la base $\mathcal{B}$, est la *matrice*

*de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$*. Si $X$ et $X'$ sont les matrices colonnes de $\mathbf{x} \in E$ dans les bases respectives $\mathcal{B}, \mathcal{B}'$, on a la formule de changement de bases

$$X = PX'. \tag{2.1}$$

S'il est besoin de préciser, on note $P_{\mathcal{B},\mathcal{B}'}$ au lieu de $P$.

### Matrice d'un système de vecteurs.

Soient $\mathbf{x_1}, \dots, \mathbf{x_r}$ un système de $r$ vecteurs de $E$. Si $X_1, \dots, X_r$ désignent les matrices colonnes des composantes de ces vecteurs dans la base $\mathcal{B} = (\mathbf{f_1}, \dots, \mathbf{f_n})$ la matrice $V = (X_1, \dots, X_r)$ est appelée matrice du système $\mathbf{x_1}, \dots, \mathbf{x_r}$ dans la base $\mathcal{B}$. C'est une matrice à $n$ lignes et $r$ colonnes. Si $\mathcal{B}' = (\mathbf{f_1'}, \dots, \mathbf{f_n'})$ est une autre base de $E$, les matrices colonnes $X_1', \dots, X_r'$ des composantes des vecteurs dans la base $\mathcal{B}'$ vérifient d'après (2.1) $X_i = PX_i'\ \forall i$, avec $P = P_{\mathcal{B},\mathcal{B}'}$. La matrice $V'$ du système $\mathbf{x_1}, \dots, \mathbf{x_r}$ dans la base $\mathcal{B}'$, vérifie donc

$$V = PV'. \tag{2.2}$$

Si $r = n$, le déterminant de la matrice $V$ est appelé *déterminant dans la base* $\mathcal{B}$ du système de vecteurs $\mathbf{x_1}, \dots, \mathbf{x_n}$ et noté $\det_{\mathcal{B}}(\mathbf{x_1}, \dots, \mathbf{x_n})$. D'après (2.2), on a

$$\det_{\mathcal{B}}(\mathbf{x_1}, \dots, \mathbf{x_n}) = \det P \det_{\mathcal{B}'}(\mathbf{x_1}, \dots, \mathbf{x_n}). \tag{2.3}$$

### Matrice d'un endomorphisme.

Un endomorphisme de $E$ est une application linéaire $f : E \to E$. Les endomorphismes de $E$ forment un $K$-espace vectoriel noté $\mathcal{L}(E)$. La matrice $A = \mathcal{M}(f, \mathcal{B}, \mathcal{D})$ d'un endomorphisme $f$ de $E$ dans deux bases $\mathcal{B} = (\mathbf{f_1}, \dots, \mathbf{f_n})$, $\mathcal{D} = (\mathbf{g_1}, \dots, \mathbf{g_n})$ est $A = \begin{pmatrix} a_1^1 & \dots & a_n^1 \\ \vdots & \dots & \vdots \\ a_1^n & \dots & a_n^n \end{pmatrix}$, où $\begin{pmatrix} a_j^1 \\ \vdots \\ a_j^n \end{pmatrix}$ est la matrice colonne des composantes du vecteur $f(\mathbf{f_j})$ dans la base $\mathcal{D}$. On dit que $\mathcal{B}$ est la *base de départ* et $\mathcal{D}$ la *base d'arrivée*. Lorsque $\mathcal{D} = \mathcal{B}$, on dit simplement que $A$ est la matrice de $f$ dans la base $\mathcal{B}$ et on note $A = \mathcal{M}(f, \mathcal{B})$.

Pour $\mathbf{x} \in E$, si $X$ est la matrice colonne des composantes de $\mathbf{x}$ dans la base $\mathcal{B}$, $Y$ la matrice colonne des composantes de $\mathbf{y} = f(\mathbf{x})$ dans la base $\mathcal{D}$, et $A = \mathcal{M}(f, \mathcal{B}, \mathcal{D})$, on a la relation :

$$Y = AX. \tag{2.4}$$

On notera que la matrice de passage $P$ de la base $\mathcal{B}$ à la base $\mathcal{B}'$ est la matrice $\mathcal{M}(Id_E, \mathcal{B}', \mathcal{B})$, où $Id_E$ désigne l'application identique de $E$ sur $E$. L'équation (2.4) donne dans ce cas l'équation (2.1).

Les bases de départ et d'arrivée $\mathcal{B}$ et $\mathcal{D}$ étant fixées, l'application

$$F_{\mathcal{B},\mathcal{D}} : \mathcal{L}(E) \to M_n(K) \tag{2.5}$$

définie par

$$f \mapsto F_{\mathcal{B},\mathcal{D}}(f) = \mathcal{M}(f, \mathcal{B}, \mathcal{D})$$

est un isomorphisme de l'espace vectoriel $\mathcal{L}(E)$ sur l'espace vectoriel $M_n(K)$ des matrices $n \times n$ à coefficients dans $K$. Notons que $\mathcal{L}(E)$ est aussi un anneau unitaire lorqu'on le munit de l'addition et de la loi $\circ$ . L'application $F_{\mathcal{B},\mathcal{D}}$ est alors aussi un isomorphisme d'anneaux.

**Groupe $GL(E)$.**

Il est immédiat que l'ensemble $GL(E)$ des automorphismes de l'espace vectoriel $E$, *i.e.* l'ensemble des bijections linéaires de $E$ sur lui-même, est un groupe pour la loi $\circ$ et que l'application (2.5) donne par restriction un isomorphisme du groupe $GL(E)$ sur le groupe $GL(n, K)$ des matrices inversibles de $M_n(K)$.

**Formule du changement de bases.**

Si $\mathcal{B}, \mathcal{B}', \mathcal{D}, \mathcal{D}'$ sont 4 bases de $E$ et si $P$ désigne la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ et $Q$ la matrice de passage de $\mathcal{D}$ à $\mathcal{D}'$, on a la formule du changement de bases :

$$B = Q^{-1}AP \tag{2.6}$$

avec $A = \mathcal{M}(f, \mathcal{B}, \mathcal{D})$ et $B = \mathcal{M}(f, \mathcal{B}', \mathcal{D}')$.

Dans le cas où les bases de départ et d'arrivée sont les mêmes, *i.e.* $\mathcal{D} = \mathcal{B}$ et $\mathcal{D}' = \mathcal{B}'$, la formule du changement de bases s'écrit simplement

$$B = P^{-1}AP \tag{2.7}$$

avec $A = \mathcal{M}(f, \mathcal{B})$, $B = \mathcal{M}(f, \mathcal{B}')$ et $P$ la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$.

**Matrices équivalentes, matrices semblables.**

Deux matrices $A, B \in M_n(K)$ sont dites semblables sur $K$ s'il existe une matrice inversible $P \in GL(n, K)$ telle que $B = P^{-1}AP$. D'après la formule du changement de bases (2.7), deux matrices sont semblables sur $K$ si et seulement si il existe un endomorphisme $f$ de $E$ et deux bases $\mathcal{B}$ et $\mathcal{B}'$ de $E$ telles que $A = \mathcal{M}(f, \mathcal{B})$ et $B = \mathcal{M}(f, \mathcal{B}')$.

Deux matrices $A, B \in M_n(K)$ sont dites équivalentes sur $K$ s'il existe deux matrices inversibles $P, Q \in GL(n, K)$ telle que $B = Q^{-1}AP$. D'après la formule du changement de bases (2.6), deux matrices sont équivalentes sur $K$ si et seulement si il existe un endomorphisme $f$ de $E$ et quatre bases $\mathcal{B}, \mathcal{D}, \mathcal{B}', \mathcal{D}'$ de $E$ telles que $A = \mathcal{M}(f, \mathcal{B}, \mathcal{D})$ et $B = \mathcal{M}(f, \mathcal{B}', \mathcal{D}')$.

La similitude sur $K$ et l'équivalence sur $K$ sont deux relations d'équivalence sur $M_n(K)$.

**Rang d'une matrice.**

Le rang d'un endomorphisme $f$ de $E$ est $r = \dim(\operatorname{Im} f)$. Le rang d'une matrice $A \in M_n(K)$ est le rang du système de ses vecteurs colonnes dans $K^n$, *i.e.* la dimension du sous-espace vectoriel de $K^n$ engendré par les vecteurs colonnes. Une matrice $A \in M_n(K)$ est de rang $r$ si et seulement si elle est équivalente sur $K$ à la matrice

$$J_r = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix} \tag{2.8}$$

avec $r$ fois 1.

$\mathcal{B}$ et $\mathcal{D}$ étant deux bases de $E$, le rang de $f$ est égal au rang de la matrice $A = \mathcal{M}(f, \mathcal{B}, \mathcal{D})$.

**Déterminant et trace.**

Soit $f$ un endomorphisme de $E$. D'après (2.7), si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases de $E$ et si $A = \mathcal{M}(f, \mathcal{B})$ et $B = \mathcal{M}(f, \mathcal{B}')$, on a $\det B = \det A$. Ainsi $\det \mathcal{M}(f, \mathcal{B})$ ne dépend pas de la base $\mathcal{B}$ mais seulement de $f$. On le note $\det f$. On a $\det(f \circ g) = (\det f)\ (\det g)\quad \forall f, g \in \mathcal{L}(E)$, et $f$ est bijective si et seulement si $\det f \neq 0$.

De même la trace $\operatorname{Tr} A = \sum_{i=1}^{n} a_i^i$ ne dépend pas de la base $\mathcal{B}$ mais seulement de $f$. On a en effet $\operatorname{Tr}(MN) = \operatorname{Tr}(NM)\ \forall M, N \in M_n(K)$, donc

$$\operatorname{Tr} B = \operatorname{Tr}(P^{-1}AP) = \operatorname{Tr}(PP^{-1}A) = \operatorname{Tr} A.$$

On la note $\operatorname{Tr} f$. On a $\operatorname{Tr}(f \circ g) = \operatorname{Tr}(g \circ f)\quad \forall f, g \in \mathcal{L}(E)$.

**Orientation d'un espace vectoriel réel.**

On suppose $K = \mathbb{R}$. Si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases de $E$, la matrice de passage $P = P_{\mathcal{B},\mathcal{B}'}$ est telle que $\det P \neq 0$, donc on a $\det P > 0$ ou $\det P < 0$. La relation

$$\mathcal{B}\, \mathcal{R}\, \mathcal{B}' \Leftrightarrow \det P_{\mathcal{B},\mathcal{B}'} > 0$$

est une relation d'équivalence dans l'ensemble de toutes les bases de $E$. Il y a 2 classes d'équivalence. On appelle *orientation* de $E$ l'une quelconque de ces classes. Une orientation étant fixée, une base appartenant à cette classe est appelée *directe*. Une base appartenant à l'autre classe est dite *rétrograde*. Le choix d'une base de $E$ détermine une orientation de $E$ pour laquelle cette base est directe.

### 2.1.2 Produit scalaire.

**Définition 2. 1.** *Soit $E$ un $\mathbb{R}$-espace vectoriel. On appelle produit scalaire sur $E$ une application*

$$(\mathbf{x}, \mathbf{y}) \mapsto (\mathbf{x}|\mathbf{y}) \tag{2.9}$$

*de $E \times E \to \mathbb{R}$ ayant les propriétés suivantes.*

*(i) L'application 2.9 est bilinéaire.*

*(ii) L'application 2.9 est symétrique, i.e.*

$$(\mathbf{x}|\mathbf{y}) = (\mathbf{y}|\mathbf{x})\quad \forall \mathbf{x}, \mathbf{y} \in E.$$

*(iii) L'application 2.9 est définie positive, i.e.*

$$(\mathbf{x}|\mathbf{x}) \geqslant 0\quad \forall \mathbf{x} \in E;$$

$$\forall \mathbf{x} \in E,\quad (\mathbf{x}|\mathbf{x}) = 0 \Rightarrow \mathbf{x} = 0.$$

Rappelons la définition d'une norme :

**Définition 2. 2.** *Soit $K$ un corps commutatif et $E$ un $K$-espace vectoriel, on appelle norme sur $E$ une application $\mathbf{x} \mapsto \|\mathbf{x}\|$ de $E$ dans $[0, +\infty[$ ayant les propriétés suivantes.*

*(i)* $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = 0$.

*(ii)* $\|\lambda\mathbf{x}\| = |\lambda|\,\|\mathbf{x}\| \quad \forall\lambda \in K, \mathbf{x} \in E$.

*(iii)* $\|\mathbf{x} + \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\| \quad \forall\mathbf{x}, \mathbf{y} \in E$ *(Inégalité triangulaire).*

La *norme associée* au produit scalaire est définie sur $E$ par

$$\|\mathbf{x}\| = \sqrt{(\mathbf{x}|\mathbf{x})}.$$

Notons aussi que l'on a :

$$|(\mathbf{x}|\mathbf{y})| \leqslant \|\mathbf{x}\|\|\mathbf{y}\| \quad \forall\mathbf{x}, \mathbf{y} \in E \quad \textit{(Inégalité de Cauchy-Schwarz)}; \tag{2.10}$$

pour tout $\mathbf{y} \in E$,

$$(\mathbf{x}|\mathbf{y}) = 0 \quad \forall\mathbf{x} \in E \quad \Rightarrow \quad \mathbf{y} = 0; \tag{2.11}$$

pour tous $\mathbf{y}, \mathbf{z} \in E$,

$$(\mathbf{x}|\mathbf{y}) = (\mathbf{x}|\mathbf{z}) \quad \forall\mathbf{x} \in E \quad \Rightarrow \quad \mathbf{y} = \mathbf{z}; \tag{2.12}$$

et enfin *l'identité de polarisation*

$$(\mathbf{x}|\mathbf{y}) = \frac{1}{2}(\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2) \quad \forall\mathbf{x}, \mathbf{y} \in E. \tag{2.13}$$

**Espace euclidien.**

**Définition 2. 3.** *On appelle espace euclidien un $\mathbb{R}$-espace vectoriel de dimension finie muni d'un produit scalaire.*

**Espace euclidien $\mathbb{R}^n$.**

Le *produit scalaire canonique sur $E = \mathbb{R}^n$* est le produit scalaire défini par

$$(\mathbf{x}|\mathbf{y}) = \sum_{j=1}^{n} x^j y^j$$

pour $\mathbf{x} = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}$, $\mathbf{y} = \begin{pmatrix} y^1 \\ \vdots \\ y^n \end{pmatrix}$. Cela s'écrit encore

$$(\mathbf{x}|\mathbf{y}) = {}^t\mathbf{x}\,\mathbf{y}$$

en identifiant les vecteurs $\mathbf{x}, \mathbf{y}$ à leurs matrices colonnes dans la base canonique $(\mathbf{e_1}, \dots, \mathbf{e_n})$ de $E$. (Si $A$ est une matrice $A = (a_i^j)$ sa transposée est la matrice ${}^tA = (a_j^i)$).

**Matrice du produit scalaire.**

Soit $E$ un espace euclidien et $\mathcal{B} = (\mathbf{f_1}, \dots, \mathbf{f_n})$ une base de $E$. On a :

$$(\mathbf{x}|\mathbf{y}) = (\sum_{i=1}^{n} \xi^i \mathbf{f_i} | \sum_{j=1}^{n} \eta^j \mathbf{f_j}) = \sum_{i,j=1}^{n} \xi^i \eta^j (\mathbf{f_i}|\mathbf{f_j}) = {}^tX\,AY$$

où $X = \begin{pmatrix} \xi^1 \\ \vdots \\ \xi^n \end{pmatrix}$, $Y = \begin{pmatrix} \eta^1 \\ \vdots \\ \eta^n \end{pmatrix}$ sont les matrices colonnes de $\mathbf{x}, \mathbf{y}$ dans la base $\mathcal{B}$ et $A$ est la matrice dont le terme de la ligne $i$ et colonne $j$ est le produit scalaire $(\mathbf{f_i}|\mathbf{f_j})$. On dit que $A$ est la matrice du produit scalaire dans la base $\mathcal{B}$. La base $\mathcal{B}$ est dite *orthonormée* si $A = I$, i.e.

$$(\mathbf{f_i}|\mathbf{f_j}) = \delta_{i,j} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j. \end{cases}$$

On démontre que tout espace euclidien possède des bases orthonormées.

**Exemple.** La base canonique de l'espace euclidien $\mathbb{R}^n$ est orthonormée. On l'appelle base orthonormée canonique.

**Isomorphisme avec le dual.**

Le dual $E^*$ d'un $K$-espace vectoriel $E$ est le $K$-espace vectoriel des formes linéaires sur $E$, *i.e.* des applications linéaires de $E$ dans $K$.

Il n'y a en général pas d'identification canonique, *i.e.* indépendante du choix d'une base de $E$, entre $E$ et $E^*$. Le produit scalaire d'un espace euclidien $E$ permet une telle identification canonique entre $E$ et $E^*$. Plus précisément :

**Proposition 2. 1.** *Soit $E$ un espace euclidien et $E^*$ le dual de $E$. L'application $\varphi$ de $E$ dans $E^*$ définie par*

$$\varphi(\mathbf{x}) = (\cdot|\mathbf{x})$$

*est un isomorphisme d'espaces vectoriels.*

**Démonstration.**

Il est immédiat que $\varphi$ est une application de $E$ dans $E^*$ et qu'elle est linéaire. Elle est injective puisque la condition $(\mathbf{y}|\mathbf{x}) = 0 \quad \forall \mathbf{y} \in E$ implique $\mathbf{x} = 0$. Comme $\dim E = \dim E^*$, elle est bijective, donc c'est un isomorphisme. □

**Supplémentaire orthogonal.**

**Proposition 2. 2.** *Soit $E$ un espace euclidien et $F$ un sous-espace vectoriel de $E$. Soit $F^\perp = \{y \in F;\ (x|y) = 0 \quad \forall x \in F\}$. Alors $F^\perp$ est un sous-espace vectoriel de $E$ et $E = F \oplus F^\perp$.*

**Démonstration.**

Il est immédiat que $F^\perp$ est un sous-espace vectoriel de $E$. On a $F \cap F^\perp = \{0\}$ puisque $\mathbf{x} \in F \cap F^\perp$ implique $(\mathbf{x}|\mathbf{x}) = 0$. La somme $F + F^\perp$ est donc directe, *i.e.* $F + F^\perp = F \oplus F^\perp$. Pour montrer que $E = F \oplus F^\perp$, considérons l'application $\psi$ de

$E$ dans le dual $F^*$ de $F$ définie par $\psi(\mathbf{x}) = (\cdot|\mathbf{x})|_F$, *i.e.* $\psi(\mathbf{x})(\mathbf{y}) = (\mathbf{y}|\mathbf{x}) \quad \forall \mathbf{y} \in F$. $\psi$ est linéaire et son noyau est $F^\perp$. Par décomposition canonique de $\psi$, on obtient donc une application linéaire injective de $E/F^\perp$ dans $F^*$. Cela implique

$$\dim(E/F^\perp) = \dim E - \dim F^\perp \leqslant \dim F^* = \dim F,$$

donc $\dim E \leqslant \dim F + \dim F^\perp = \dim\left(F \oplus F^\perp\right)$ et ainsi $E = F \oplus F^\perp$. □

**Définition 2. 4.** *On appelle $F^\perp$ l'orthogonal ou encore le supplémentaire orthogonal de $F$.*

### 2.1.3 Adjoint d'un endomorphisme.

**Théorème 2. 1 (adjoint d'un endomorphisme).** *Soit $E$ un espace euclidien et $f$ un endomorphisme de $E$.*
*(i) Il existe un endomorphisme $f^*$ de $E$ unique tel que*

$$(f(\mathbf{x})|\mathbf{y}) = (\mathbf{x}|f^*(\mathbf{y})) \quad \forall \mathbf{x}, \mathbf{y} \in E. \tag{2.14}$$

*(ii) Dans toute base orthonormée, la matrice de $f^*$ est la transposée de celle de $f$.*

**Démonstration.**
(i) Si $f^*$ existe, son unicité est immédiate, puisque s'il existait deux telles applications $f^*$ et $f^\dagger$ on aurait

$$(f(\mathbf{x})|\mathbf{y}) = (\mathbf{x}|f^*(\mathbf{y})) = (\mathbf{x}|f^\dagger(\mathbf{y})) \quad \forall \mathbf{x}, \mathbf{y} \in E,$$

donc, d'après l'implication (2.12) :

$$f^\dagger(\mathbf{y}) = f^*(\mathbf{y}) \quad \forall \mathbf{y} \in E.$$

Montrons l'existence de $f^*$. Soit $\mathbf{y} \in E$. L'application $\psi_\mathbf{y}$ de $E$ dans $K$ définie par $\psi_\mathbf{y}(\mathbf{x}) = (f(\mathbf{x})|\mathbf{y})$ $\forall \mathbf{x} \in E$ est linéaire, donc est un élément du dual $E^*$. D'après l'identification canonique entre $E$ et $E^*$, il existe donc un unique vecteur $\mathbf{v_y} \in E$ tel que $\psi_\mathbf{y}(\mathbf{x}) = (\mathbf{x}|\mathbf{v_y})$ $\forall \mathbf{x} \in E$, *i.e.*

$$(f(\mathbf{x})|\mathbf{y}) = (\mathbf{x}|\mathbf{v_y}) \quad \forall \mathbf{x} \in E.$$

On a pour tous $\mathbf{x}, \mathbf{y}, \mathbf{y}' \in E$

$$\begin{aligned}(\mathbf{x}|\mathbf{v}_{\mathbf{y}+\mathbf{y}'}) &= (f(\mathbf{x})|\mathbf{y}+\mathbf{y}') \\ &= (f(\mathbf{x})|\mathbf{y}) + (f(\mathbf{x})|\mathbf{y}') \\ &= (\mathbf{x}|\mathbf{v_y}) + (\mathbf{x}|\mathbf{v}_{\mathbf{y}'}) \\ &= (\mathbf{x}|\mathbf{v_y} + \mathbf{v}_{\mathbf{y}'})\end{aligned}$$

donc, d'après l'implication (2.12) : $\mathbf{v}_{\mathbf{y}+\mathbf{y}'} = \mathbf{v_y} + \mathbf{v}_{\mathbf{y}'}$. De même, pour $\lambda \in K$,

$$(\mathbf{x}|\mathbf{v}_{\lambda\mathbf{y}}) = (f(\mathbf{x})|\lambda\mathbf{y}) = \lambda(f(\mathbf{x})|\mathbf{y}) = \lambda(\mathbf{x}|\mathbf{v_y}) = (\mathbf{x}|\lambda\mathbf{v_y})$$

donc $\mathbf{v}_{\lambda\mathbf{y}} = \lambda\mathbf{v_y}$. L'application $f^* : E \to E$ définie par $f^*(\mathbf{y}) = \mathbf{v_y}$ $\forall \mathbf{y} \in E$ est donc un endomorphisme de $E$.

(ii) Si maintenant $\mathcal{B} = (\mathbf{f_1}, \dots, \mathbf{f_n})$ est une base orthonormée quelconque et $A = (a^i_j)$ désigne la matrice de $f$ dans la base $\mathcal{B}$, on a :

$$(\mathbf{f_i}|f^*(\mathbf{f_j})) = (f(\mathbf{f_i})|\mathbf{f_j}) = (\sum_{k=1}^{n} a^k_i \mathbf{f_k}|\mathbf{f_j}) = \sum_{k=1}^{n} a^k_i (\mathbf{f_k}|\mathbf{f_j}) = a^j_i$$

donc $\mathcal{M}(f^*, \mathcal{B}) = {}^t A$. □

**Définition 2. 5.** *Soit $E$ un espace euclidien et $f$ un endomorphisme de $E$. L'endomorphisme $f^*$ du Théorème 2.1 est appelé l'endomorphisme adjoint de $f$.*

**Proposition 2. 3.** *Soit $E$ un espace euclidien. On a :*
*(i)* $(Id_E)^* = Id_E$.
*(ii)* $(f^*)^* = f^* \quad \forall f \in \mathcal{L}(E)$.
*(iii)* $(f \circ g)^* = g^* \circ f^* \quad \forall f, g \in \mathcal{L}(E)$.
*(iv) Si $f \in \mathcal{L}(E)$ est bijective, il en est de même de $f^*$ et $(f^*)^{-1} = (f^{-1})^*$.*
*(v)* $\det f^* = \det f \quad \forall f \in \mathcal{L}(E)$.

**Démonstration.**
Fixons une base orthonormée $\mathcal{B}$ de $E$. On sait que pour tout endomorphisme $f$ de $E$, la matrice de $f^*$ dans la base $\mathcal{B}$ est la transposée de celle de $f$. En considérant les matrices, les propriétés (i) à (v) sont alors immédiates. □

## 2.2 Isométries, groupe orthogonal.

Dans toute la suite, $E$ désigne un espace euclidien.

### 2.2.1 Endomorphismes isométriques.

**Proposition 2. 4.** *Soit $f$ un endomorphisme de $E$. Les propriétés suivantes sont équivalentes :*
*(i)* $\|f(\mathbf{x})\| = \|\mathbf{x}\| \quad \forall \mathbf{x} \in E$.
*(ii)* $(f(\mathbf{x})|f(\mathbf{y})) = (\mathbf{x}|\mathbf{y}) \quad \forall \mathbf{x}, \mathbf{y} \in E$.
*(iii)* $f^* \circ f = Id_E$.
*(iv) $f$ est bijective et $f^* = f^{-1}$.*

**Démonstration.**
(i) ⇒ (ii). On a pour tous $\mathbf{x}, \mathbf{y} \in E$ d'après l'identité de polarisation (2.13) :

$$\begin{aligned}(f(\mathbf{x})|f(\mathbf{y})) &= \frac{1}{2}(\|f(\mathbf{x}) + f(\mathbf{y})\|^2 - \|f(\mathbf{x})\|^2 - \|f(\mathbf{y})\|^2) \\ &= \frac{1}{2}(\|f(\mathbf{x} + \mathbf{y})\|^2 - \|f(\mathbf{x})\|^2 - \|f(\mathbf{y})\|^2) \\ &= \frac{1}{2}(\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2) \\ &= (\mathbf{x}|\mathbf{y}).\end{aligned}$$

(ii) ⇒ (iii). (ii) s'écrit $(\mathbf{x}|f^*(f(\mathbf{y}))) = (\mathbf{x}|\mathbf{y}) \; \forall \mathbf{x}, \mathbf{y} \in E$ d'où $f^*(f(\mathbf{y})) = \mathbf{y} \; \forall \mathbf{y} \in E$ et $f^* \circ f = Id_E$.
(iii) ⇒ (iv). $\det f^* \det f = \det(f^* \circ f) = \det Id_E = 1$, donc $\det f \neq 0$ et $f$ est

bijective. On obtient alors $f^* = f^{-1}$ en composant à droite l'égalité (iii) par $f^{-1}$.
(iv) $\Rightarrow$ (i). $\|f(\mathbf{x})\|^2 = (f(\mathbf{x})|f(\mathbf{x})) = (\mathbf{x}|(f^*(f(\mathbf{x}))) = (\mathbf{x}|\mathbf{x}) = \|\mathbf{x}\|^2$ puisque $f^* \circ f = Id_E$. □

**Définition 2. 6.** *On dit qu'un endomorphisme de $E$ est isométrique s'il possède l'une des propriétés équivalentes de la proposition 2.4.*

**Proposition 2. 5.** *(i) Un endomorphisme isométrique de $E$ a pour déterminant* $\pm 1$.
*(ii) L'ensemble noté $O(E)$ des endomorphismes isométriques de $E$ est un groupe pour la loi* $\circ$.
*(iii) Les endomorphismes isométriques de déterminant* 1 *forment un sous-groupe distingué de $O(E)$ noté $SO(E)$.*

**Démonstration.**
(i) On sait que $\det f^* = \det f \quad \forall f \in \mathcal{L}(E)$. Or un endomorphisme est isométrique si et seulement si $f^* \circ f = Id_E$. (Prop. 2.4). Donc si $f$ est isométrique, $\det(f^* \circ f) = (\det f)^2 = 1$.
(ii) D'après (i), $O(E)$ est un sous-ensemble du groupe $GL(E)$ des automorphismes de l'espace vectoriel $E$. C'est un sous-groupe de $GL(E)$ puisque :
• $Id_E \in O(E)$.
• $\forall f, g \in O(E), \forall \mathbf{x} \in E, \|(g \circ f)(\mathbf{x})\| = \|g(f(\mathbf{x}))\| = \|f(\mathbf{x})\| = \|\mathbf{x}\|$ donc $g \circ f \in O(E)$.
• $\forall f \in O(E)$, on a $f^{-1} \in O(E)$ puisque pour tout $\mathbf{y} \in E$, on a en posant $\mathbf{x} = f^{-1}(\mathbf{y})$ : $\|f^{-1}(\mathbf{y})\| = \|\mathbf{x}\| = \|f(\mathbf{x})\| = \|\mathbf{y}\|$.
(iii) $SO(E)$ est un sous-ensemble de $O(E)$.
• $Id_E \in SO(E)$.
• $\forall f, g \in SO(E), g \circ f \in SO(E)$ puisque $\det(g \circ f) = (\det g)(\det f) = 1$.
• $\forall f \in SO(E)$, on a $f^{-1} \in SO(E)$ puisque $\det f^{-1} = (\det f)^{-1} = 1$.
$SO(E)$ est donc un sous-groupe de $O(E)$. Il est distingué dans $O(E)$ puisque pour tout $f \in SO(E)$ et tout $g \in O(E)$, $g \circ f \circ g^{-1} \in O(E)$ et $\det(g \circ f \circ g^{-1}) = \det f = 1$ donc $g \circ f \circ g^{-1} \in SO(E)$. □

**Définition 2. 7.** *Le groupe $O(E)$ des endomorphismes isométriques de $E$ est appelé le groupe orthogonal de $E$. Le sous-groupe $SO(E)$ est appelé le groupe des rotations de $E$.*

## 2.2.2 Matrices orthogonales.

**Définition 2. 8.** *Une matrice $A \in M_n(\mathbb{R})$ est dite orthogonale si elle vérifie :*

$$ {}^tA\,A = A\,{}^tA = I \tag{2.15} $$

**Théorème 2. 2.** *Soit $f$ un endomorphisme de $E$, $\mathcal{B} = (\mathbf{f_1}, \dots, \mathbf{f_n})$ une base orthonormée de $E$, et $A = \mathcal{M}(f, \mathcal{B})$ la matrice de $f$ dans la base $\mathcal{B}$. Les deux propriétés suivantes sont équivalentes.*
*(i) L'endomorphisme $f$ est isométrique.*
*(ii) La matrice $A$ est orthogonale.*

**Démonstration.**
On a :

$$\begin{aligned} f \text{ isométrique} &\Leftrightarrow f^* \circ f = Id_E && \text{(Prop. 2.4)} \\ &\Leftrightarrow {}^tA\,A = I && \text{(Th. 2.1)} \\ &\Leftrightarrow A \text{ inversible et } {}^tA = A^{-1} \\ &\Leftrightarrow A \text{ orthogonale.} \end{aligned}$$

□

D'après (2.15), si $A$ est une matrice orthogonale réelle, on a

$$\det {}^tA\,A = (\det A)^2 = \det I = 1,$$

donc $\det A = \pm 1$.

On note $O(n,\mathbb{R})$ ou simplement $O(n)$ l'ensemble des matrices orthogonales réelles, et $SO(n,\mathbb{R})$ ou $SO(n)$ l'ensemble des matrices orthogonales réelles de déterminant 1.

**Théorème 2. 3.** *(i) $O(n)$ est un sous-groupe de $GL(n,\mathbb{R})$ isomorphe à $O(E)$.*
*(ii) $SO(n)$ est un sous-groupe distingué de $O(n)$ isomorphe à $SO(E)$.*

**Démonstration.**
(i) Fixons une base orthonormée $\mathcal{B} = (\mathbf{f_1}, \ldots, \mathbf{f_n})$ de $E$, et considérons l'application $F : GL(E) \to GL(n,\mathbb{R})$ qui associe à un endomorphisme $f$ de $E$ sa matrice $A = \mathcal{M}(f,\mathcal{B})$ dans la base $\mathcal{B}$. $F$ est un isomorphisme du groupe $GL(E)$ sur le groupe $GL(n,\mathbb{R})$. L'image par $F$ du sous-groupe $O(E)$ de $GL(E)$ est $O(n)$ d'après le Théorème 2.2. Donc $O(n)$ est un sous-groupe de $GL(n,\mathbb{R})$, et la restriction de $F$ à $O(E)$ est un isomorphisme de $O(E)$ sur $O(n)$.
(ii) L'image par $F$ du sous-groupe $SO(E)$ de $GL(E)$ est $SO(n)$. Donc $SO(n)$ est un sous-groupe de $GL(n,\mathbb{R})$ et la restriction de $F$ à $SO(E)$ est un isomorphisme de $SO(E)$ sur $SO(n)$. Comme $SO(E)$ est distingué dans $O(E)$, le sous-groupe image $F(SO(E)) = SO(n)$ est distingué dans $F(O(E)) = O(n)$ (Th. 1.6).

□

**Proposition 2. 6.** *Une matrice $A \in M_n(\mathbb{R})$ est orthogonale si et seulement si ses vecteurs-colonnes $\{\mathbf{c_1}, \ldots, \mathbf{c_n}\}$ forment un système orthonormé de $\mathbb{R}^n$ pour le produit scalaire canonique.*

**Démonstration.**
Si $A = (a^i_j)$, on a :

$$\begin{aligned} A \in O(n) &\Leftrightarrow {}^tA\,A = I \\ &\Leftrightarrow \sum_{k=1}^{n} a^k_i a^k_j = \delta_{i,j} \quad \forall i,j \\ &\Leftrightarrow (\mathbf{c_i}|\mathbf{c_j}) = \delta_{i,j} \quad \forall i,j. \end{aligned}$$

□

**Corollaire.** *Soit $A \in M_n(\mathbb{R})$ et $\mathcal{B} = (\mathbf{f_1}, \ldots, \mathbf{f_n})$ une base orthonormée de $E$. $A$ est orthogonale si et seulement si c'est la matrice de passage de la base $\mathcal{B}$ à une autre base orthonormée $\mathcal{B}' = (\mathbf{f'_1}, \ldots, \mathbf{f'_n})$ de $E$.*

**Démonstration.**

Observons d'abord que la colonne d'indice $j$ ($1 \leqslant j \leqslant n$) de $A$ est la colonne des composantes dans la base $\mathcal{B}$ du vecteur

$$\mathbf{f}'_j = \sum_{i=1}^{n} a^i_j \mathbf{f}_i \in E.$$

Si $\mathcal{B}' = (\mathbf{f}'_1, \dots, \mathbf{f}'_n)$ est une base de $E$, on a donc $A = P_{\mathcal{B},\mathcal{B}'}$ par définition de la matrice de passage.

Or

$$(\mathbf{f}'_i|\mathbf{f}'_j) = \sum_{k=1}^{n} a^k_i a^k_j = (\mathbf{c}_i|\mathbf{c}_j) \quad \forall i,j \quad (1 \leqslant i,j \leqslant n), \tag{2.16}$$

donc $\mathcal{B}' = (\mathbf{f}'_1, \dots, \mathbf{f}'_n)$ est une base orthonormée de $E$ si et seulement si les vecteurs colonnes $\{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ de $A$ forment un système orthonormé de $\mathbb{R}^n$ pour le produit scalaire canonique, *i.e.* si et seulement si $A$ est une matrice orthogonale.

Ainsi, si $A$ est orthogonale, $\mathcal{B}'$ est une base orthonormée de $E$ et alors $A = P_{\mathcal{B},\mathcal{B}'}$. Réciproquement, si $\mathcal{B}'$ est une base orthonormée de $E$, $A = P_{\mathcal{B},\mathcal{B}'}$ et la matrice $A$ est orthogonale d'après (2.16). □

### 2.2.3 Produit vectoriel en dimension 3.

On suppose l'espace vectoriel euclidien $E$ orienté et de dimension 3. Si $\mathcal{B}, \mathcal{B}'$ sont deux bases orthonormées directes de $E$, on a pour $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \in E$

$$\det_{\mathcal{B}}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \det_{\mathcal{B}'}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3).$$

En effet, d'après (2.3)

$$\det_{\mathcal{B}}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \det P \det_{\mathcal{B}'}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$$

avec $P = P_{\mathcal{B},\mathcal{B}'}$. Or $P$ est orthogonale puisque $\mathcal{B}$ et $\mathcal{B}'$ sont orthonormées; son déterminant est donc $\pm 1$. C'est 1 puisque $\mathcal{B}$ et $\mathcal{B}'$ sont directes.

On appelle *produit mixte* des trois vecteurs $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ et on note $[\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3]$ le déterminant $\det_{\mathcal{B}}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ dans une base orthonormée directe quelconque de $E$. Comme on vient de le voir, il ne dépend pas de la base orthonormée directe de $E$ utilisée.

D'après les propriétés du déterminant, $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ est une base (*resp.* directe) de $E$ si et seulement si $[\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3] \neq 0$ (*resp.* $> 0$).

Fixons maintenant $\mathbf{x}_1, \mathbf{x}_2 \in E$ et considérons l'application $E$ dans $\mathbb{R}$ définie par $\mathbf{x} \mapsto [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}]$. C'est une forme linéaire sur $E$. D'après la Prop.2.1, il existe un vecteur $\mathbf{w} \in E$ unique tel que

$$(\mathbf{x}|\mathbf{w}) = [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}] \quad \forall \mathbf{x} \in E.$$

On dit que $\mathbf{w}$ est le *produit vectoriel* de $\mathbf{x}_1$ et $\mathbf{x}_2$ et on note $\mathbf{w} = \mathbf{x}_1 \wedge \mathbf{x}_2$. On a donc par définition

$$[\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}] = (\mathbf{x}|\mathbf{x}_1 \wedge \mathbf{x}_2) \quad \forall \mathbf{x} \in E.$$

**Proposition 2. 7.** *(i) Soient $\mathbf{x}_1, \mathbf{x}_2 \in E$. Le produit vectoriel $\mathbf{x}_1 \wedge \mathbf{x}_2$ est nul si et seulement si $\mathbf{x}_1$ et $\mathbf{x}_2$ sont liés.*

*(ii)* $\mathbf{x_1} \wedge \mathbf{x_2}$ *est orthogonal à* $\mathbf{x_1}$ *et à* $\mathbf{x_2}$.
*(iii) Si* $\mathcal{B} = (\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3})$ *est une base orthonormée directe de* $E$ *et pour* $i = 1, 2$ $\mathbf{x_i} = \sum_{j=1}^{3} \xi_i^j \mathbf{e_j}$, $(\xi_i^j \in \mathbb{R})$, *alors* $\mathbf{x_1} \wedge \mathbf{x_2} = \sum_{j=1}^{3} \eta^j \mathbf{e_j}$ *avec*

$$\begin{aligned} \eta^1 &= \xi_1^2 \xi_2^3 - \xi_1^3 \xi_2^2 \\ \eta^2 &= \xi_1^3 \xi_2^1 - \xi_1^1 \xi_2^3 \\ \eta^3 &= \xi_1^1 \xi_2^2 - \xi_1^2 \xi_2^1 . \end{aligned}$$

*(iv) On a*

$$\|\mathbf{x_1} \wedge \mathbf{x_2}\|^2 = \|\mathbf{x_1}\|^2 \, \|\mathbf{x_2}\|^2 - (\mathbf{x_1}|\mathbf{x_2})^2 \quad (\textit{identité de Lagrange}).$$

*(v) Si* $\mathbf{x_1}$ *et* $\mathbf{x_2}$ *sont indépendants,* $(\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_1} \wedge \mathbf{x_2})$ *est une base directe de* $E$.
*(vi) Si* $\mathbf{x_1}$ *et* $\mathbf{x_2}$ *sont orthogonaux et normés,* $(\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_1} \wedge \mathbf{x_2})$ *est une base orthonormée directe de* $E$.
*(vii) L'application* $(\mathbf{x_1}, \mathbf{x_2}) \mapsto \mathbf{x_1} \wedge \mathbf{x_2}$ *de* $E \times E$ *dans* $E$ *est bilinéaire. Elle est antisymétrique, i.e.*

$$\mathbf{x_2} \wedge \mathbf{x_1} = - \, \mathbf{x_1} \wedge \mathbf{x_2} \quad \forall \mathbf{x_1}, \mathbf{x_2} \in E.$$

*(viii) Pour* $\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3} \in E$,

$$\mathbf{x_1} \wedge (\mathbf{x_2} \wedge \mathbf{x_3}) = (\mathbf{x_1}|\mathbf{x_3})\, \mathbf{x_2} - (\mathbf{x_1}|\mathbf{x_2})\, \mathbf{x_3} \tag{2.17}$$
$$(\mathbf{x_1} \wedge \mathbf{x_2}) \wedge \mathbf{x_3} = (\mathbf{x_1}|\mathbf{x_3})\, \mathbf{x_2} - (\mathbf{x_2}|\mathbf{x_3})\, \mathbf{x_1} \tag{2.18}$$

*En particulier la loi définie par* $(\mathbf{x_1}, \mathbf{x_2}) \mapsto \mathbf{x_1} \wedge \mathbf{x_2}$ *dans* $E$ *n'est pas associative.*
*(ix) Pour* $\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3} \in E$, *on a*

$$\mathbf{x_1} \wedge (\mathbf{x_2} \wedge \mathbf{x_3}) + \mathbf{x_2} \wedge (\mathbf{x_3} \wedge \mathbf{x_1}) + \mathbf{x_3} \wedge (\mathbf{x_1} \wedge \mathbf{x_2}) = 0 \quad (\textit{identité de Jacobi}).$$

**Démonstration.**
(i) Si $\mathbf{x_1}$ et $\mathbf{x_2}$ sont liés, on a $[\mathbf{x_1}, \mathbf{x_2}, \mathbf{x}] = 0$ pour tout $\mathbf{x} \in E$. En particulier pour $\mathbf{x} = \mathbf{x_1} \wedge \mathbf{x_2}$, cela donne

$$0 = [\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_1} \wedge \mathbf{x_2}] = (\mathbf{x_1} \wedge \mathbf{x_2} \,|\, \mathbf{x_1} \wedge \mathbf{x_2}) = \|\mathbf{x_1} \wedge \mathbf{x_2}\|^2$$

donc $\mathbf{x_1} \wedge \mathbf{x_2} = 0$. Montrons que réciproquement la condition $\mathbf{x_1} \wedge \mathbf{x_2} = 0$ implique que $\mathbf{x_1}$ et $\mathbf{x_2}$ sont liés. Par contraposition, cela équivaut à montrer que la condition $\mathbf{x_1}$ et $\mathbf{x_2}$ indépendants implique $\mathbf{x_1} \wedge \mathbf{x_2} \neq 0$. Or si $\mathbf{x_1}$ et $\mathbf{x_2}$ sont indépendants, il existe $\mathbf{x} \in E$ tel que $(\mathbf{x_1}, \mathbf{x_2}, \mathbf{x})$ soit une base de $E$. Alors $[\mathbf{x_1}, \mathbf{x_2}, \mathbf{x}] \neq 0$. Mais $[\mathbf{x_1}, \mathbf{x_2}, \mathbf{x}] = (\mathbf{x}|\mathbf{x_1} \wedge \mathbf{x_2})$, donc $\mathbf{x_1} \wedge \mathbf{x_2} \neq 0$.
(ii) $(\mathbf{x_1}|\mathbf{x_1} \wedge \mathbf{x_2}) = [\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_1}] = 0$, et de même $(\mathbf{x_2}|\mathbf{x_1} \wedge \mathbf{x_2}) = [\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_2}] = 0$.
(iii) On a

$$\eta^1 = (\mathbf{e_1}|\mathbf{x_1} \wedge \mathbf{x_2}) = [\mathbf{x_1}, \mathbf{x_2}, \mathbf{e_1}] = \begin{vmatrix} \xi_1^1 & \xi_2^1 & 1 \\ \xi_1^2 & \xi_2^2 & 0 \\ \xi_1^3 & \xi_2^3 & 0 \end{vmatrix} = \xi_1^2 \xi_2^3 - \xi_1^3 \xi_2^2,$$

$$\eta^2 = (\mathbf{e_2}|\mathbf{x_1}\wedge\mathbf{x_2}) = [\mathbf{x_1},\mathbf{x_2},\mathbf{e_2}] = \begin{vmatrix} \xi_1^1 & \xi_2^1 & 0 \\ \xi_1^2 & \xi_2^2 & 1 \\ \xi_1^3 & \xi_2^3 & 0 \end{vmatrix} = \xi_1^3\xi_2^1 - \xi_1^1\xi_2^3,$$

$$\eta^3 = (\mathbf{e_3}|\mathbf{x_1}\wedge\mathbf{x_2}) = [\mathbf{x_1},\mathbf{x_2},\mathbf{e_3}] = \begin{vmatrix} \xi_1^1 & \xi_2^1 & 0 \\ \xi_1^2 & \xi_2^2 & 0 \\ \xi_1^3 & \xi_2^3 & 1 \end{vmatrix} = \xi_1^1\xi_2^2 - \xi_1^2\xi_2^1.$$

(iv)

$$\begin{aligned}
\|\mathbf{x_1}\|^2\,\|\mathbf{x_2}\|^2 - (\mathbf{x_1}|\mathbf{x_2})^2 &= ((\xi_1^1)^2 + (\xi_1^2)^2 + (\xi_1^3)^2)\,((\xi_2^1)^2 + (\xi_2^2)^2 + (\xi_2^3)^2) \\
&\quad -(\xi_1^1\xi_2^1 + \xi_1^2\xi_2^2 + \xi_1^3\xi_2^3)^2 \\
&= \underline{(\xi_1^1)^2(\xi_2^1)^2} + (\xi_1^1)^2(\xi_2^2)^2 + (\xi_1^1)^2(\xi_2^3)^2 - \underline{(\xi_1^1)^2(\xi_2^1)^2} \\
&\quad +(\xi_1^2)^2(\xi_2^1)^2 + \underbrace{(\xi_1^2)^2(\xi_2^2)^2} + (\xi_1^2)^2(\xi_2^3)^2 - \underbrace{(\xi_1^2)^2(\xi_2^2)^2} \\
&\quad +(\xi_1^3)^2(\xi_2^1)^2 + (\xi_1^3)^2(\xi_2^2)^2 + \overbrace{(\xi_1^3)^2(\xi_2^3)^2} - \overbrace{(\xi_1^3)^2(\xi_2^3)^2} \\
&\quad -2\xi_1^1\xi_1^2\xi_2^1\xi_2^2 - 2\xi_1^1\xi_1^3\xi_2^1\xi_2^3 - 2\xi_1^2\xi_1^3\xi_2^2\xi_2^3 \\
&= (\xi_1^1\xi_2^2 - \xi_1^2\xi_2^1)^2 + (\xi_1^3\xi_2^1 - \xi_1^1\xi_2^3)^2 + (\xi_1^2\xi_2^3 - \xi_1^3\xi_2^2)^2 \\
&= (\eta^3)^2 + (\eta^2)^2 + (\eta^1)^2 \\
&= \|\mathbf{x_1}\wedge\mathbf{x_2}\|^2 .
\end{aligned}$$

(v) D'après (i), $\mathbf{x_1}\wedge\mathbf{x_2} \neq 0$ puisque $\mathbf{x_1}$ et $\mathbf{x_2}$ sont indépendants. Maintenant, $[\mathbf{x_1},\mathbf{x_2},\mathbf{x_1}\wedge\mathbf{x_2}] = \|\mathbf{x_1}\wedge\mathbf{x_2}\|^2 > 0$ donc $(\mathbf{x_1},\mathbf{x_2},\mathbf{x_1}\wedge\mathbf{x_2})$ est une base de $E$ et elle est directe.

(vi) Le système $\{\mathbf{x_1},\mathbf{x_2}\}$ étant orthonormé, $\mathbf{x_1}$ et $\mathbf{x_2}$ sont indépendants. D'après (v) $(\mathbf{x_1},\mathbf{x_2},\mathbf{x_1}\wedge\mathbf{x_2})$ est une base directe de $E$. Il reste à voir qu'elle est orthonormée. Or d'après (ii), $\mathbf{x_1}\wedge\mathbf{x_2}$ est orthogonal à $\mathbf{x_1}$ et $\mathbf{x_2}$, et d'après l'identité de Lagrange, $\|\mathbf{x_1}\wedge\mathbf{x_2}\| = 1$.

(vii) Pour tout $\mathbf{x}\in E$ on a d'après les propriétés du déterminant

$$[\mathbf{x_2},\mathbf{x_1},\mathbf{x}] = -\,[\mathbf{x_1},\mathbf{x_2},\mathbf{x}]\,.$$

Cela s'écrit

$$(\mathbf{x}|\mathbf{x_2}\wedge\mathbf{x_1}) = -(\mathbf{x}|\mathbf{x_1}\wedge\mathbf{x_2}) \quad \forall\,\mathbf{x}\in E.$$

D'après (2.12), cela équivaut à $\mathbf{x_2}\wedge\mathbf{x_1} = -\,\mathbf{x_1}\wedge\mathbf{x_2}$. Donc l'application $(\mathbf{x_1},\mathbf{x_2}) \mapsto \mathbf{x_1}\wedge\mathbf{x_2}$ de $E\times E$ dans $E$ est antisymétrique. Montrons qu'elle est bilinéaire. Comme elle est antisymétrique, il suffit de démontrer que

$$(\lambda\mathbf{x_1} + \mu\mathbf{y_1})\wedge\mathbf{x_2} = \lambda(\mathbf{x_1}\wedge\mathbf{x_2}) + \mu(\mathbf{y_1}\wedge\mathbf{x_2})$$

pour tous $\lambda,\mu\in\mathbb{R}$ et tous $\mathbf{x_1},\mathbf{y_1},\mathbf{x_2}\in E$. Or pour tout $\mathbf{x}\in E$, on a d'après les propriétés du déterminant

$$\begin{aligned}
(\mathbf{x}\,|\,(\lambda\mathbf{x_1}+\mu\mathbf{y_1})\wedge\mathbf{x_2}) &= [\lambda\mathbf{x_1}+\mu\mathbf{y_1},\mathbf{x_2},\mathbf{x}] \\
&= \lambda\,[\mathbf{x_1},\mathbf{x_2},\mathbf{x}] + \mu\,[\mathbf{y_1},\mathbf{x_2},\mathbf{x}] \\
&= \lambda(\mathbf{x}\,|\,\mathbf{x_1}\wedge\mathbf{x_2}) + \mu(\mathbf{x}\,|\,\mathbf{y_1}\wedge\mathbf{x_2}) \\
&= (\mathbf{x}\,|\,\lambda(\mathbf{x_1}\wedge\mathbf{x_2}) + \mu(\mathbf{y_1}\wedge\mathbf{x_2}))
\end{aligned}$$

d'où le résultat d'après (2.12).
(viii) Fixons une base orthonormée directe $\mathcal{B} = (\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3})$ de $E$. En identifiant un vecteur à la matrice-colonne de ses composantes dans la base $\mathcal{B}$, on a $\mathbf{x_1} = \begin{pmatrix} \xi_1^1 \\ \xi_1^2 \\ \xi_1^3 \end{pmatrix}$, $\mathbf{x_2} = \begin{pmatrix} \xi_2^1 \\ \xi_2^2 \\ \xi_2^3 \end{pmatrix}$, $\mathbf{x_3} = \begin{pmatrix} \xi_3^1 \\ \xi_3^2 \\ \xi_3^3 \end{pmatrix}$, donc

$$
\begin{aligned}
\mathbf{x_1} \wedge (\mathbf{x_2} \wedge \mathbf{x_3}) &= \begin{pmatrix} \xi_1^1 \\ \xi_1^2 \\ \xi_1^3 \end{pmatrix} \wedge \left( \begin{pmatrix} \xi_2^1 \\ \xi_2^2 \\ \xi_2^3 \end{pmatrix} \wedge \begin{pmatrix} \xi_3^1 \\ \xi_3^2 \\ \xi_3^3 \end{pmatrix} \right) \\
&= \begin{pmatrix} \xi_1^1 \\ \xi_1^2 \\ \xi_1^3 \end{pmatrix} \wedge \begin{pmatrix} \xi_2^2\xi_3^3 - \xi_2^3\xi_3^2 \\ \xi_2^3\xi_3^1 - \xi_2^1\xi_3^3 \\ \xi_2^1\xi_3^2 - \xi_2^2\xi_3^1 \end{pmatrix} \\
&= \begin{pmatrix} \xi_1^2(\xi_2^1\xi_3^2 - \xi_2^2\xi_3^1) - \xi_1^3(\xi_2^3\xi_3^1 - \xi_2^1\xi_3^3) \\ \xi_1^3(\xi_2^2\xi_3^3 - \xi_2^3\xi_3^2) - \xi_1^1(\xi_2^1\xi_3^2 - \xi_2^2\xi_3^1) \\ \xi_1^1(\xi_2^3\xi_3^1 - \xi_2^1\xi_3^3) - \xi_1^2(\xi_2^2\xi_3^3 - \xi_2^3\xi_3^2) \end{pmatrix} \\
&= \begin{pmatrix} \xi_2^1(\xi_1^2\xi_3^2 + \xi_1^3\xi_3^3) - \xi_3^1(\xi_1^2\xi_2^2 + \xi_1^3\xi_2^3) \\ \xi_2^2(\xi_1^1\xi_3^1 + \xi_1^3\xi_3^3) - \xi_3^2(\xi_1^1\xi_2^1 + \xi_1^3\xi_2^3) \\ \xi_2^3(\xi_1^1\xi_3^1 + \xi_1^2\xi_3^2) - \xi_3^3(\xi_1^1\xi_2^1 + \xi_1^2\xi_2^2) \end{pmatrix} \\
&= \begin{pmatrix} \xi_2^1(\xi_1^1\xi_3^1 + \xi_1^2\xi_3^2 + \xi_1^3\xi_3^3) - \xi_3^1(\xi_1^1\xi_2^1 + \xi_1^2\xi_2^2 + \xi_1^3\xi_2^3) \\ \xi_2^2(\xi_1^1\xi_3^1 + \xi_1^2\xi_3^2 + \xi_1^3\xi_3^3) - \xi_3^2(\xi_1^1\xi_2^1 + \xi_1^2\xi_2^2 + \xi_1^3\xi_2^3) \\ \xi_2^3(\xi_1^1\xi_3^1 + \xi_1^2\xi_3^2 + \xi_1^3\xi_3^3) - \xi_3^3(\xi_1^1\xi_2^1 + \xi_1^2\xi_2^2 + \xi_1^3\xi_2^3) \end{pmatrix} \\
&= (\xi_1^1\xi_3^1 + \xi_1^2\xi_3^2 + \xi_1^3\xi_3^3) \begin{pmatrix} \xi_2^1 \\ \xi_2^2 \\ \xi_2^3 \end{pmatrix} - (\xi_1^1\xi_2^1 + \xi_1^2\xi_2^2 + \xi_1^3\xi_2^3) \begin{pmatrix} \xi_3^1 \\ \xi_3^2 \\ \xi_3^3 \end{pmatrix} \\
&= (\mathbf{x_1}|\mathbf{x_3})\,\mathbf{x_2} - (\mathbf{x_1}|\mathbf{x_2})\,\mathbf{x_3}\,.
\end{aligned}
$$

Cela montre (2.17). Pour (2.18), on utilise l'antisymétrie:

$$(\mathbf{x_1} \wedge \mathbf{x_2}) \wedge \mathbf{x_3} = -\,\mathbf{x_3} \wedge (\mathbf{x_1} \wedge \mathbf{x_2}) = -(\mathbf{x_3}|\mathbf{x_2})\,\mathbf{x_1} + (\mathbf{x_3}|\mathbf{x_1})\,\mathbf{x_2}.$$

(ix)

$$
\begin{aligned}
&\mathbf{x_1} \wedge (\mathbf{x_2} \wedge \mathbf{x_3}) + \mathbf{x_2} \wedge (\mathbf{x_3} \wedge \mathbf{x_1}) + \mathbf{x_3} \wedge (\mathbf{x_1} \wedge \mathbf{x_2}) \\
&\quad = (\mathbf{x_1}|\mathbf{x_3})\,\mathbf{x_2} - (\mathbf{x_1}|\mathbf{x_2})\,\mathbf{x_3} + (\mathbf{x_2}|\mathbf{x_1})\,\mathbf{x_3} - (\mathbf{x_2}|\mathbf{x_3})\,\mathbf{x_1} + (\mathbf{x_3}|\mathbf{x_2})\,\mathbf{x_1} - (\mathbf{x_3}|\mathbf{x_1})\,\mathbf{x_2} \\
&\quad = 0.
\end{aligned}
$$

□

### 2.2.4 Semi-simplicité des endomorphismes isométriques.

**Théorème 2. 4.** *Soit $f$ un endomorphisme isométrique de $E$ et $F$ un sous-espace vectoriel stable par $f$, i.e. $f(F) \subset F$.*

*Alors l'orthogonal $F^\perp$ de $F$ est aussi stable par $f$.*

**Démonstration.**

On sait que $E = F \oplus F^{\perp}$. Soit $\mathbf{y} \in F^{\perp}$. Pour tout $\mathbf{x} \in F$, on a :

$$(f(\mathbf{y})|\mathbf{x}) = (\mathbf{y}|f^{*}(\mathbf{x})) = (\mathbf{y}|f^{-1}(\mathbf{x})).$$

Comme $f$ est bijective on a $\dim f(F) = \dim F$, donc $f(F) \subset F$ implique $f(F) = F$ et par conséquent $f^{-1}(\mathbf{x}) \in F$.

Alors $(f(\mathbf{y})|\mathbf{x}) = (\mathbf{y}|f^{-1}(\mathbf{x})) = 0$. Comme $\mathbf{x} \in F$ est arbitraire, $f(\mathbf{y}) \in F^{\perp}$. Comme $\mathbf{y} \in F^{\perp}$ est arbitraire, $F^{\perp}$ est stable par $f$.

### 2.2.5 Calcul de $SO(2)$ et $O(2)$.

On utilise dans ce qui suit les fonctions sin et cos ainsi que la notion d'angle de deux vecteurs non nuls de l'espace euclidien $\mathbb{R}^2$. Ces notions feront l'objet d'une introduction rigoureuse indépendante au chapitre 7.

**Théorème 2. 5.** *(i)*

$$\begin{aligned} SO(2) &= \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} ;\ a, b \in \mathbb{R},\ a^2 + b^2 = 1 \right\} \\ &= \left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} ;\ \theta \in \mathbb{R} \right\}. \end{aligned}$$

*(ii) Le groupe $SO(2)$ est commutatif.*

*(iii)*

$$\begin{aligned} O(2) \setminus SO(2) &= \left\{ \begin{pmatrix} a & b \\ b & -a \end{pmatrix} ;\ a, b \in \mathbb{R},\ a^2 + b^2 = 1 \right\} \\ &= \left\{ \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} ;\ \theta \in \mathbb{R} \right\}. \end{aligned}$$

**Démonstration.**

(i) Par définition,

$$SO(2) = \{A \in M_2(\mathbb{R});\ {}^tA\,A = I;\ det\,A = 1\}.$$

Soit donc $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R})$. On a

$$A \in SO(2) \Leftrightarrow \det A = 1 \text{ et } {}^tA = A^{-1}.$$

Or on sait que $A$ est inversible si et seulement si $\det A \neq 0$, et que l'on a alors

$$A^{-1} = (\det A)^{-1} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Les conditions

$$\begin{aligned} &\det A = 1 \\ &{}^tA = A^{-1} \end{aligned}$$

s'écrivent donc

$$\begin{aligned} ad - bc &= 1 \\ \begin{pmatrix} a & c \\ b & d \end{pmatrix} &= \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \end{aligned}$$

ce qui est équivalent à

$$\begin{aligned} ad - bc &= 1 \\ \begin{pmatrix} a & c \\ b & d \end{pmatrix} &= \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \end{aligned}$$

ou encore

$$\begin{aligned} d &= a \\ c &= -b \\ a^2 + b^2 &= 1\,. \end{aligned}$$

Les éléments de $SO(2)$ sont donc les matrices $A$ de la forme

$$A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix};\ a, b \in \mathbb{R},\ a^2 + b^2 = 1.$$

La deuxième égalité de (i) résulte du fait que pour tout couple $(a, b)$ de réels tels que $a^2 + b^2 = 1$ il existe $\theta \in \mathbb{R}$ tel que $a = \cos\theta$ et $b = -\sin\theta$.
(ii) Pour $A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ et $B = \begin{pmatrix} c & d \\ -d & c \end{pmatrix}$ éléments de $SO(2)$ $(a, b, c, d \in \mathbb{R},\ a^2 + b^2 = 1,\ c^2 + d^2 = 1)$, on a

$$AB = \begin{pmatrix} ac - bd & ad + bc \\ -(ad + bc) & ac - bd \end{pmatrix} = BA.$$

(iii) La matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R})$ appartient à $O(2) \setminus SO(2)$ si et seulement si $A \in O(2)$ et $\det A = -1$. Les conditions

$$\begin{aligned} \det A &= -1 \\ {}^tA &= A^{-1} \end{aligned}$$

s'écrivent

$$\begin{aligned} ad - bc &= -1 \\ \begin{pmatrix} a & c \\ b & d \end{pmatrix} &= \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \end{aligned}$$

ce qui est équivalent à

$$\begin{aligned} ad - bc &= -1 \\ \begin{pmatrix} a & c \\ b & d \end{pmatrix} &= -\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \end{aligned}$$

ou encore

$$\begin{aligned} d &= -a \\ c &= b \\ a^2+b^2 &= 1\,. \end{aligned}$$

Les éléments de $O(2) \setminus SO(2)$ sont donc les matrices $A$ de la forme

$$A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix};\ a, b \in \mathbb{R},\ a^2+b^2 = 1.$$

La deuxième égalité de (iii) résulte du fait que pour tout couple $(a,b)$ de réels tels que $a^2+b^2=1$ il existe $\theta \in \mathbb{R}$ tel que $a = \cos\theta$ et $b = \sin\theta$.

On aurait aussi pu utiliser le raisonnement qui suit. Soit $J = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Il est immédiat que $J \in O(2) \setminus SO(2)$ puisque $\det J = -1$, $J^2 = I$ et $J$ est symétrique. Considérons l'application $L_J : O(2) \to O(2)$ définie par $A \mapsto L_J(A) = JA$. Tout $B \in O(2)$ possède comme antécédent unique par $L_J$ la matrice $A = JB \in O(2)$. L'application $L_J$ est donc une bijection de $O(2)$ sur lui-même. Comme $\det B = \det J \det A = -\det A$, on a $B \in O(2) \setminus SO(2) \Leftrightarrow A \in SO(2)$. L'image par $L_J$ de $SO(2)$ est donc $O(2) \setminus SO(2)$. Cela s'écrit encore

$$O(2) \setminus SO(2) = L_J(SO(2)) = J\,SO(2).$$

Or

$$J \begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}.$$

Donc d'après (i), on retrouve que les éléments de $O(2) \setminus SO(2)$ sont les matrices $A$ de la forme

$$A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix};\ a, b \in \mathbb{R},\ a^2+b^2 = 1.$$

□

**Remarque.** Soit $\theta \in \mathbb{R}$. L'endomorphisme de $\mathbb{R}^2$ dont la matrice dans la base orthonormée canonique est

$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

sera appelé *rotation (vectorielle) d'angle* $\theta$ au Chap. 7. L'endomorphisme de $\mathbb{R}^2$ dont la matrice dans la base orthonormée canonique est

$$\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$$

est la symétrie orthogonale par rapport à la droite de pente $\tan\frac{\theta}{2}$ dont un vecteur directeur est $\mathbf{v} = \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} \end{pmatrix}$ (Exercice 2.2.).

### 2.2.6 Réduction canonique d'un élément de $SO(3)$.

Dans toute cette section, $E$ désigne un espace vectoriel euclidien de dimension 3, muni d'une base orthonormée $(\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3})$ fixée (par exemple la base orthonormée canonique si $E = \mathbb{R}^n$), et de l'orientation définie par cette base.

**Rotation d'axe orienté $\mathbb{R}\mathbf{k}$ et d'angle $\theta$.**

**Proposition 2. 8.** *Soit* $\mathbf{k}$ *un vecteur normé de* $E$ *et* $\theta \in \mathbb{R}$. *Soit* $(\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3)$ *une base orthonormée directe de* $E$ *telle que* $\mathbf{k} = \mathbf{f}_3$, *et* $R_{\mathbf{k}}(\theta) \in \mathcal{L}(E)$ *l'endomorphisme de* $E$ *dont la matrice dans la base* $(\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3)$ *est*

$$B = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

*Alors :*
*(i)* $R_{\mathbf{k}}(\theta) \in SO(E)$.
*(ii) La matrice de* $R_{\mathbf{k}}(\theta)$ *dans toute base orthonormée directe* $(\mathbf{f}'_1, \mathbf{f}'_2, \mathbf{f}'_3)$ *telle que* $\mathbf{f}'_3 = \mathbf{k}$ *est encore* $B$.
*(iii) Si* $\theta \not\equiv 0 \pmod{2\pi}$ *l'ensemble des vecteurs* $\mathbf{x} \in E$ *invariants par* $R_{\mathbf{k}}(\theta)$ *est la droite* $\Delta = \mathbb{R}\mathbf{k}$.
*(iv) Si* $\theta \not\equiv 0 \pmod{2\pi}$, *l'équation* $R_{\mathbf{k}}(\theta) = R_{\mathbf{k}'}(\theta')$ *(*$\mathbf{k}'$ *vecteur normé,* $\theta' \in \mathbb{R}$*) a lieu si et seulement si* $\mathbf{k}' = \mathbf{k}$ *et* $\theta' \equiv \theta \pmod{2\pi}$ *ou si* $\mathbf{k}' = -\mathbf{k}$ *et* $\theta' \equiv -\theta \pmod{2\pi}$.

**Démonstration.**
(i) On vérifie facilement que

$${}^tB\,B = B\,{}^tB = I$$

et $\det B = 1$. Donc $B \in SO(3)$ et $R_{\mathbf{k}}(\theta) \in SO(E)$.
(ii) Soit $(\mathbf{f}'_1, \mathbf{f}'_2, \mathbf{f}'_3)$ une autre base orthonormée directe telle que $\mathbf{f}'_3 = \mathbf{k}$. La matrice de passage $P$ de la base $(\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3)$ à la base $(\mathbf{f}'_1, \mathbf{f}'_2, \mathbf{f}'_3)$ appartient à $SO(3)$, car les deux bases sont orthonormées directes. Mais elle est de la forme

$$\begin{pmatrix} \multicolumn{2}{c}{\multirow{2}{*}{$Q$}} & 0 \\ & & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

et l'on a $P \in SO(3) \Leftrightarrow Q \in SO(2)$. Maintenant, la matrice de $R_{\mathbf{k}}(\theta)$ dans la base $(\mathbf{f}'_1, \mathbf{f}'_2, \mathbf{f}'_3)$ est

$$\begin{aligned} P^{-1}BP &= \begin{pmatrix} Q^{-1} & 0 \\ & 0 \\ 0 \quad 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} Q & 0 \\ & 0 \\ 0 \quad 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} S & 0 \\ & 0 \\ 0 \quad 0 & 1 \end{pmatrix} \end{aligned}$$

où

$$S = Q^{-1} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} Q.$$

Mais $SO(2)$ est un groupe commutatif (Th. 2.5), donc

$$S = Q^{-1}Q \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

et alors

$$P^{-1}BP = B.$$

(iii) Le polynôme caractéristique de $R_{\mathbf{k}}(\theta)$ est

$$\det(R_{\mathbf{k}}(\theta) - X\,Id_E) = \det(B - XI) = (X^2 - 2X\cos\theta + 1)(1 - X).$$

Ses racines sont $1, e^{i\theta} = \cos\theta + i\sin\theta$ et $e^{-i\theta} = \cos\theta - i\sin\theta$. Si $\theta \not\equiv 0 \pmod{2\pi}$, $e^{i\theta}$ et $e^{-i\theta}$ sont différents de 1, donc 1 est une valeur propre *simple* de $R_{\mathbf{k}}(\theta)$. Or l'ensemble des vecteurs invariants par $R_{\mathbf{k}}(\theta)$ est le sous-espace propre $\mathrm{Ker}\,(R_{\mathbf{k}}(\theta) - Id_E)$. Il est donc de dimension 1 puisque 1 est une valeur propre simple. Il contient $\mathbf{k}$, donc c'est $\mathbb{R}\mathbf{k}$.
(iv) Supposons $R_{\mathbf{k}}(\theta) = R_{\mathbf{k}'}(\theta')$ avec $\mathbf{k}'$ vecteur normé et $\theta' \in \mathbb{R}$. Comme $\mathbf{k}'$ est invariant par $R_{\mathbf{k}'}(\theta')$, il est invariant par $R_{\mathbf{k}}(\theta)$, donc c'est un multiple de $\mathbf{k}$ d'après (iii). Soit $\lambda \in \mathbb{R}$ tel que $\mathbf{k}' = \lambda\mathbf{k}$. Comme $\mathbf{k}$ et $\mathbf{k}'$ sont normés, $1 = \|\mathbf{k}'\| = |\lambda|\|\mathbf{k}\| = |\lambda|$ donc $\lambda = \pm 1$. Si $\lambda = 1, \mathbf{k}' = \mathbf{k}$. En comparant les matrices de $R_{\mathbf{k}}(\theta)$ et $R_{\mathbf{k}}(\theta')$ dans la base $(\mathbf{f_1}, \mathbf{f_2}, \mathbf{f_3})$ on voit alors que l'équation $R_{\mathbf{k}}(\theta) = R_{\mathbf{k}}(\theta')$ implique $\theta' \equiv \theta \pmod{2\pi}$. Si $\lambda = -1, \mathbf{k}' = -\mathbf{k}$. Posons alors $\mathbf{f}'_1 = \mathbf{f_2}, \mathbf{f}'_2 = \mathbf{f_1}, \mathbf{f}'_3 = \mathbf{k}' = -\mathbf{k}$. $(\mathbf{f}'_1, \mathbf{f}'_2, \mathbf{f}'_3)$ est une base orthonormée directe telle que $\mathbf{f}'_3 = \mathbf{k}'$ donc la matrice de $R_{\mathbf{k}'}(\theta')$ dans cette base est

$$\begin{pmatrix} \cos\theta' & -\sin\theta' & 0 \\ \sin\theta' & \cos\theta' & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Or il est immédiat que la matrice de $R_{\mathbf{k}}(\theta)$ dans cette base est

$$\begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Ces deux matrices sont égales si et seulement si $\theta' \equiv -\theta \pmod{2\pi}$.

Ce dernier point montre aussi que $R_{-\mathbf{k}}(-\theta) = R_{\mathbf{k}}(\theta)$, et la réciproque en résulte alors. □

**Définition 2. 9.** *L'endomorphisme $R_{\mathbf{k}}(\theta) \in \mathcal{L}(E)$ de la Proposition 2.8 est appelé rotation d'axe orienté $\mathbb{R}\mathbf{k}$ et d'angle $\theta$. Identifiant un endomorphisme de $E$ à sa matrice dans la base orthonormée fixée $(\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3})$, on notera aussi $R_{\mathbf{k}}(\theta) \in SO(3)$ la matrice dans la base $(\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3})$ de l'endomorphisme $R_{\mathbf{k}}(\theta)$.*

**Proposition 2. 9.** *Pour tout vecteur normé $\mathbf{k}$, on a :*
*(i)* $R_{\mathbf{k}}(\theta + \varphi) = R_{\mathbf{k}}(\theta)R_{\mathbf{k}}(\varphi) \quad \forall\theta, \varphi \in \mathbb{R}$.
*(ii)* $Tr\,R_{\mathbf{k}}(\theta) = 1 + 2\cos\theta$.
*(iii) Pour tout vecteur $\mathbf{v}$ non colinéaire avec $\mathbf{k}$, $\sin\theta$ est du signe du produit mixte* $[\mathbf{v}, R_{\mathbf{k}}(\theta)\mathbf{v}, \mathbf{k}]$.

**Démonstration.**

(i) Soit $(\mathbf{f_1}, \mathbf{f_2}, \mathbf{f_3})$ une base orthonormée de $E$ telle que $\mathbf{k} = \mathbf{f_3}$, et $P$ la matrice de passage de la base orthonormée canonique $(\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3})$ à la base $(\mathbf{f_1}, \mathbf{f_2}, \mathbf{f_3})$. On a pour tout $t \in \mathbb{R}$

$$R_{\mathbf{k}}(t) = P \begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix} P^{-1},$$

donc pour tous $\theta, \varphi \in \mathbb{R}$

$$\begin{aligned}
R_{\mathbf{k}}(\theta)R_{\mathbf{k}}(\varphi) &= P \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} P^{-1} P \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} P^{-1} \\
&= P \begin{pmatrix} \cos(\theta+\varphi) & -\sin(\theta+\varphi) & 0 \\ \sin(\theta+\varphi) & \cos(\theta+\varphi) & 0 \\ 0 & 0 & 1 \end{pmatrix} P^{-1} \\
&= R_{\mathbf{k}}(\theta+\varphi).
\end{aligned}$$

(ii) Résulte immédiatement de la définition de $R_{\mathbf{k}}(\theta)$.
(iii) On sait que le produit mixte d'un système de vecteurs ne dépend pas de la base orthonormée directe dans laquelle on le calcule. On se place donc dans la base $(\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3)$. Soit $X = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$ la matrice-colonne des composantes de $\mathbf{v}$ dans la base $(\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3)$. La matrice-colonne des composantes de $R_{\mathbf{k}}(\theta)\mathbf{v}$ est alors $Y = \begin{pmatrix} \alpha\cos\theta - \beta\sin\theta \\ \alpha\sin\theta + \beta\cos\theta \\ \gamma \end{pmatrix}$, donc

$$[\mathbf{v}, R_{\mathbf{k}}(\theta)\mathbf{v}, \mathbf{k}] = \begin{vmatrix} \alpha & \alpha\cos\theta - \beta\sin\theta & 0 \\ \beta & \alpha\sin\theta + \beta\cos\theta & 0 \\ \gamma & \gamma & 1 \end{vmatrix} = (\alpha^2 + \beta^2)\sin\theta\,.$$

D'où le résultat puisque $\alpha^2 + \beta^2 > 0$. □

**Réduction canonique d'un élément de $SO(3)$.**

**Théorème 2. 6.** *Soit $A \in SO(3), A \neq I$ et $f$ l'endomorphisme de $E$ dont la matrice dans la base orthonormée $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ fixée est $A$.*

*(i) Il existe une base orthonormée directe $(\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3)$ de $E$ dans laquelle la matrice de l'endomorphisme $f$ est*

$$B = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

*avec $\theta \in \mathbb{R}, \theta \not\equiv 0 \pmod{2\pi}$.*
*(ii) Si $P$ désigne la matrice de passage de la base $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ à la base $(\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3)$, $P \in SO(3)$ et $B = P^{-1}AP$.*
*(iii) On a $f = R_{\mathbf{k}}(\theta)$ avec $\mathbf{k} = \mathbf{f}_3$.*

**Démonstration.**

(i) Montrons d'abord que 1 est valeur propre de $A$. On a :

$$\begin{aligned}
\det(A - I) &= \det(A - {}^tA\,A) \\
&= \det((I - {}^tA)A) \\
&= \det(I - {}^tA)\det A \\
&= \det(I - {}^tA) \text{ (car } \det A = 1) \\
&= \det {}^t(I - A) \\
&= \det(I - A) \\
&= -\det(A - I) \text{ (car } \dim E = 3)
\end{aligned}$$

donc $\det(A - I) = 0$ et 1 est valeur propre de $A$.

Soit $\mathbf{f_3}$ un vecteur normé de $E$ tel que $A\mathbf{f_3} = \mathbf{f_3}$, et $\Delta = \mathbb{R}\mathbf{f_3}$. $\Delta$ est une droite vectorielle dont $\mathbf{f_3}$ est un vecteur directeur. L'orthogonal $\Delta^\perp$ de $\Delta$ est un plan vectoriel, et l'on a :

$$E = \Delta \oplus \Delta^\perp.$$

Soit $(\mathbf{f_1}, \mathbf{f_2})$ une base de $\Delta^\perp$ telle que la base $\mathcal{B} = (\mathbf{f_1}, \mathbf{f_2}, \mathbf{f_3})$ de $E$ soit *directe*. L'endomorphisme $f$ laisse stable $\Delta^\perp$ ( Th. 2.4), donc la matrice $B$ de $f$ dans la base $\mathcal{B}$ est de la forme

$$\begin{pmatrix} \multicolumn{2}{c}{\multirow{2}{*}{C}} & 0 \\ & & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Or $A \in SO(3) \Leftrightarrow B \in SO(3) \Leftrightarrow C \in SO(2)$. Donc il existe $\theta \in \mathbb{R}$ tel que

$$C = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Comme $A \neq I$, on a aussi $C \neq I$, et donc $\theta \not\equiv 0 \pmod{2\pi}$.
(ii) et (iii) résultent immédiatement de (i). □

### 2.2.7 Classes de conjugaison de $SO(3)$.

**Théorème 2. 7.** *Soit* $\mathbf{k}$ *un vecteur normé et* $\theta \in \mathbb{R}$. *On a pour tout* $B \in SO(3)$

$$BR_{\mathbf{k}}(\theta)B^{-1} = R_{B\mathbf{k}}(\theta). \tag{2.19}$$

**Démonstration.**

Soit $\mathcal{B} = (\mathbf{f_1}, \mathbf{f_2}, \mathbf{f_3})$ une base orthonormée directe de $E$ telle que $\mathbf{k} = \mathbf{f_3}$. La matrice de la rotation $R_{\mathbf{k}}(\theta)$ dans la base $\mathcal{B}$ est

$$\mathcal{M}(R_{\mathbf{k}}(\theta), \mathcal{B}) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Comme $B \in SO(3)$, $\mathcal{B}' = (B\mathbf{f_1}, B\mathbf{f_2}, B\mathbf{f_3})$ une base orthonormée directe de $E$. La matrice $\mathcal{M}(B^{-1}, \mathcal{B}', \mathcal{B})$ de la rotation $B^{-1}$ dans les deux bases $\mathcal{B}'$ , $\mathcal{B}$ est la matrice identité $I$. La matrice $\mathcal{M}(B, \mathcal{B}, \mathcal{B}')$ de la rotation $B$ dans les deux bases $\mathcal{B}$,$\mathcal{B}'$ est aussi la matrice identité $I$. La matrice de la rotation $BR_{\mathbf{k}}(\theta)B^{-1}$ dans la base $\mathcal{B}'$ est donc

$$\begin{aligned} \mathcal{M}(BR_{\mathbf{k}}(\theta)B^{-1}, \mathcal{B}') &= \mathcal{M}(B, \mathcal{B}, \mathcal{B}')\mathcal{M}(R_{\mathbf{k}}(\theta), \mathcal{B})\mathcal{M}(B^{-1}, \mathcal{B}', \mathcal{B}) \\ &= I\mathcal{M}(R_{\mathbf{k}}(\theta), \mathcal{B})I \\ &= \mathcal{M}(R_{\mathbf{k}}(\theta), \mathcal{B}) \\ &= \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \mathcal{M}(R_{B\mathbf{k}}(\theta), \mathcal{B}') \end{aligned}$$

donc

$$BR_{\mathbf{k}}(\theta)B^{-1} = R_{B\mathbf{k}}(\theta).$$

□

**Corollaire.** *Soit* $A \in SO(3)$. *L'inverse* $A^{-1}$ *est conjugué à* $A$ *dans* $SO(3)$, *i.e. il existe* $B \in SO(3)$ *tel que* $BAB^{-1} = A^{-1}$.

**Démonstration.**

Si $A = R_{\mathbf{k}}(\theta)$, il suffit de prendre $B = R_{\mathbf{u}}(\pi)$, avec un vecteur normé $\mathbf{u}$ orthogonal à $\mathbf{k}$. On a alors en effet $B\mathbf{k} = R_{\mathbf{u}}(\pi)\mathbf{k} = -\mathbf{k}$, donc

$$BAB^{-1} = BR_{\mathbf{k}}(\theta)B^{-1} = R_{B\mathbf{k}}(\theta) = R_{-\mathbf{k}}(\theta) = R_{\mathbf{k}}(-\theta) = A^{-1}. \tag{2.20}$$

□

**Remarque.** On notera que l'équation (2.20) s'écrit encore

$$R_{\mathbf{u}}(\pi)R_{\mathbf{k}}(\theta)R_{\mathbf{u}}(\pi) = R_{\mathbf{k}}(-\theta) \tag{2.21}$$

pour $\mathbf{k}, \mathbf{u}$ vecteurs normés orthogonaux.

**Théorème 2. 8.** *L'ensemble des classes de conjugaison de $SO(3)$ est en bijection avec $[0, \pi]$.*

**Démonstration.**

Soit $A = R_{\mathbf{k}}(\theta) \in SO(3)$ tel que $\theta \not\equiv 0 \pmod{2\pi}$. On sait que (Prop. 2.8) pour $\theta' \in \mathbb{R}$ et $\mathbf{k}'$ normé on a $A = R_{\mathbf{k}'}(\theta')$ si et seulement si $\mathbf{k}' = \mathbf{k}$ et $\theta' \equiv \theta \pmod{2\pi}$ ou si $\mathbf{k}' = -\mathbf{k}$ et $\theta' \equiv -\theta \pmod{2\pi}$.

En particulier, la condition $R_{\mathbf{k}}(\theta) = R_{\mathbf{k}'}(\theta')$ avec $\theta, \theta' \in [0, \pi]$ équivaut si $\theta \neq 0, \pi$ à $\mathbf{k}' = \mathbf{k}$ et $\theta' = \theta$. Si $\theta = \pi$, elle équivaut à l'ensemble des deux possibilités $\mathbf{k}' = \mathbf{k}$ et $\theta' = \pi$ ou $\mathbf{k}' = -\mathbf{k}$ et $\theta' = \pi$. Si $\theta = 0$, elle équivaut à $\theta' = 0$.

Dans tous les cas, on constate que $\theta' = \theta$. Soit alors $A \in SO(3)$. Il existe $\mathbf{k}$ normé et $\theta \in [0, \pi]$ tel que $A = R_{\mathbf{k}}(\theta)$. On vient de voir que si $A = R_{\mathbf{k}'}(\theta')$ avec $\theta' \in [0, \pi]$, on a nécessairement $\theta' = \theta$. Donc $\theta$ ne dépend que de $A$.

Si l'on pose $\theta = F(A)$, on définit une application $F : SO(3) \to [0, \pi]$. Cette application $F$ est surjective. Soit $A \in SO(3)$ et $\theta = F(A) \in [0, \pi]$. On a donc $A = R_{\mathbf{k}}(\theta)$ avec $\mathbf{k}$ vecteur normé. Pour tout $A' \in SO(3)$, la définition de $F$ montre que $F(A') = F(A)$ si et seulement si il existe un vecteur normé $\mathbf{k}'$ tel que $A' = R_{\mathbf{k}'}(\theta)$. Mais il existe un $B \in SO(3)$ tel que $B(\mathbf{k}) = \mathbf{k}'$. La condition équivaut donc à dire qu'il existe $B \in SO(3)$ tel que $A' = R_{B\mathbf{k}}(\theta)$. Or d'après (2.19) $R_{B\mathbf{k}}(\theta) = BAB^{-1}$. On a donc $F(A) = F(A')$ si et seulement si il existe $B \in SO(3)$ tel que $A' = BAB^{-1}$, *i.e.* $A'$ est conjugué à $A$ dans $SO(3)$.

On en déduit que $F$ induit une bijection $F_*$ de l'ensemble des classes de conjugaison de $SO(3)$ sur l'intervalle $[0, \pi]$. □

## 2.3 Groupe euclidien.

$E$ désigne un espace vectoriel euclidien.

### 2.3.1 Isométries.

**Définition 2. 10.** *On dit qu'une application $f$ de $E$ dans lui-même est une isométrie si l'on a pour tous $\mathbf{x}, \mathbf{y} \in E$ :*

$$\|f(\mathbf{x}) - f(\mathbf{y})\| = \|\mathbf{x} - \mathbf{y}\|.$$

**Théorème 2. 9.** *Une application $f$ de $E$ dans lui-même est une isométrie si et seulement si il existe un couple $(\mathbf{a}, g)$, où $\mathbf{a} \in E$ et $g$ est un endomorphisme isométrique de $E$, tel que $f = \mathbf{a} + g$, i.e.*

$$f(\mathbf{x}) = \mathbf{a} + g(\mathbf{x}) \quad \forall \mathbf{x} \in E. \tag{2.22}$$

*Un tel couple $(\mathbf{a}, g)$ est unique.*

**Démonstration.**

*Condition suffisante.* Soit $(\mathbf{a}, g)$ un couple avec $\mathbf{a} \in E$ et $g$ endomorphisme isométrique de $E$, et soit $f$ l'application de $E$ dans lui-même définie par $f(\mathbf{x}) = \mathbf{a} + g(\mathbf{x}) \quad \forall \mathbf{x} \in E$. Alors $f$ est une isométrie de $E$ puisque pour tous $\mathbf{x}, \mathbf{y} \in E$ :

$$\begin{aligned} \|f(\mathbf{x}) - f(\mathbf{y})\| &= \|g(\mathbf{x}) - g(\mathbf{y})\| \\ &= \|g(\mathbf{x} - \mathbf{y})\| \\ &= \|\mathbf{x} - \mathbf{y}\|. \end{aligned}$$

*Condition nécessaire.* Soit $f$ une isométrie de $E$ dans lui-même. Montrons qu'il existe un couple unique $(\mathbf{a}, g)$ avec $\mathbf{a} \in E$ et $g$ endomorphisme de $E$, tel que $f = \mathbf{a} + g$, et que l'endomorphisme $g$ est alors isométrique. Si $(\mathbf{a}, g)$ est un couple avec $\mathbf{a} \in E$ et $g$ endomorphisme de $E$, tel que $f = \mathbf{a} + g$, on a nécessairement $g(0) = 0$ donc $\mathbf{a} = f(0)$ et alors

$$g(\mathbf{x}) = f(\mathbf{x}) - f(0) \quad \forall \mathbf{x} \in E. \tag{2.23}$$

Cela prouve l'unicité d'un tel couple. Pour prouver son existence, définissons $\mathbf{a}$ par $\mathbf{a} = f(0)$ et l'application $g$ par la formule (2.23). On a donc $f = \mathbf{a} + g$. Il reste à montrer que $g$ est un endomorphisme isométrique. L'équation $\|f(\mathbf{x}) - f(0)\| = \|\mathbf{x} - 0\|$ s'écrit $\|g(\mathbf{x})\| = \|\mathbf{x}\| \quad \forall \mathbf{x} \in E$. Cela implique que

$$(g(\mathbf{x})|g(\mathbf{y})) = (\mathbf{x}|\mathbf{y}) \quad \forall \mathbf{x}, \mathbf{y} \in E. \tag{2.24}$$

En effet,

$$\begin{aligned} -2(g(\mathbf{x})|g(\mathbf{y})) &= \|g(\mathbf{x}) - g(\mathbf{y})\|^2 - \|g(\mathbf{x})\|^2 - \|g(\mathbf{y})\|^2 \\ &= \|f(\mathbf{x}) - f(\mathbf{y})\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2 \\ &= \|\mathbf{x} - \mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2 \\ &= -2(\mathbf{x}|\mathbf{y}). \end{aligned}$$

D'après (2.24), l'image par $g$ d'une base orthonormée est une base orthonormée. Fixons une base orthonormée $(\mathbf{e_1}, \ldots, \mathbf{e_n})$ de $E$ (par exemple la base canonique si $E = \mathbb{R}^n$). L'image $(\mathbf{g}(\mathbf{e_1}), \ldots, \mathbf{g}(\mathbf{e_n}))$ de la base orthonormée $(\mathbf{e_1}, \ldots, \mathbf{e_n})$ est une base orthonormée. Pour tout vecteur $\mathbf{y}$, on a donc

$$\mathbf{y} = \sum_{j=1}^{n} (\mathbf{y}|g(\mathbf{e_j}))g(\mathbf{e_j}).$$

En particulier,

$$\begin{aligned} g(\mathbf{x}) &= \sum_{j=1}^{n} (g(\mathbf{x})|g(\mathbf{e_j}))g(\mathbf{e_j}) \quad \forall \mathbf{x} \in E \\ &= \sum_{j=1}^{n} (\mathbf{x}|\mathbf{e_j})g(\mathbf{e_j}) \quad \forall \mathbf{x} \in E. \end{aligned}$$

D'après cette formule, l'application $g$ est linéaire. D'autre part d'après (2.24), $g$ est isométrique. □

**Corollaire.** *(i) Toute isométrie de E dans lui-même est bijective.*
*(ii) L'ensemble des isométries de E est un groupe pour la loi* $\circ$.

*(iii) L'application* $\Theta$ *de* $E \times O(E)$ *dans l'ensemble des isométries de* $E$ *définie par* $\Theta(\mathbf{a}, g) = f$ *avec*

$$f(\mathbf{x}) = \mathbf{a} + g(\mathbf{x}) \quad \forall \mathbf{x} \in E$$

*est une bijection.*

**Démonstration.**

(i) Soit $f$ une isométrie de $E$. Il existe un couple $(\mathbf{a}, g)$, où $\mathbf{a} \in E$ et $g$ est un endomorphisme isométrique de $E$, vérifiant $f = \mathbf{a} + g$. Comme $g$ est un endomorphisme isométrique de $E$, il est bijectif (Prop. 2.4). Or $f = \mathbf{a} + g$ s'écrit encore $f = t_{\mathbf{a}} \circ g$ en notant $t_{\mathbf{a}} : E \to E$ la translation $\mathbf{x} \mapsto t_{\mathbf{a}}(\mathbf{x}) = \mathbf{a} + \mathbf{x}$. Comme $t_{\mathbf{a}}$ est bijectif, $g$ est la composée de deux bijections, donc est bijective.

(ii) Il suffit de vérifier que l'ensemble des isométries de $E$ est un sous-groupe du groupe Bij (E).

- L'ensemble des isométries de $E$ est inclus dans Bij (E) d'après (i).
- $Id_E$ est clairement une isométrie de $E$.
- Si $f, g$ sont deux isométries de $E$, on a pour tous $\mathbf{x}, \mathbf{y} \in E$ :

$$\|(g \circ f)(\mathbf{x}) - (g \circ f)(\mathbf{y})\| = \|g(f(\mathbf{x})) - g(f(\mathbf{y}))\| = \|f(\mathbf{x}) - f(\mathbf{y})\| = \|\mathbf{x} - \mathbf{y}\|$$

donc $g \circ f$ est une isométrie.

- Si $f$ est une isométrie de $E$, on a pour tout $\mathbf{x}, \mathbf{y} \in E$

$$\|\mathbf{x} - \mathbf{y}\| = \|f(f^{-1}(\mathbf{x})) - f(f^{-1}(\mathbf{y}))\| = \|f^{-1}(\mathbf{x}) - f^{-1}(\mathbf{y})\|$$

donc $f^{-1}$ est une isométrie.

(iii) Résulte directement du Théorème 2.9 . □

## 2.3.2 Groupe euclidien.

**Définition 2. 11.** *Le groupe des isométries de* $E$ *est appelé le groupe euclidien de* $E$ *et est noté* $\mathcal{E}(E)$.

Fixons une base orthonormée $(\mathbf{e_1}, \ldots, \mathbf{e_n})$ de $E$ (par exemple la base canonique si $E = \mathbb{R}^n$). Cela permet d'identifier $E$ et $\mathbb{R}^n$ et également d'identifier un endomorphisme isométrique à sa matrice dans la base fixée. Alors $E \times O(E)$ est identifié à $\mathbb{R}^n \times O(n)$ et la bijection $\Theta : E \times O(E) \longrightarrow \mathcal{E}(E)$ devient : $\Theta(\mathbf{a}, A) = f$ avec

$$f(\mathbf{x}) = \mathbf{a} + A\mathbf{x} \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

**Théorème 2. 10.** *(i) Soient* $(\mathbf{a}, A), (\mathbf{b}, B) \in \mathbb{R}^n \times O(n)$. *Alors*

$$\Theta^{-1}\left[\Theta(\mathbf{a}, A) \circ \Theta(\mathbf{b}, B)\right] = (\mathbf{a} + A\mathbf{b}, AB).$$

*(ii)* $\mathbb{R}^n \times O(n)$ *muni de la loi*

$$(\mathbf{a}, A)(\mathbf{b}, B) = (\mathbf{a} + A\mathbf{b}, AB) \tag{2.25}$$

*est un groupe isomorphe au groupe* $\mathcal{E}(E)$.

**Démonstration.**

(i) On a pour tout $\mathbf{x} \in E$ :

$$\begin{aligned}(\Theta(\mathbf{a}, A) \circ \Theta(\mathbf{b}, B))(\mathbf{x}) &= \Theta(\mathbf{a}, A)(\Theta(\mathbf{b}, B)(\mathbf{x})) \\ &= \Theta(\mathbf{a}, A)(\mathbf{b} + B\mathbf{x}) \\ &= \mathbf{a} + A(\mathbf{b} + B\mathbf{x}) \\ &= \mathbf{a} + A\mathbf{b} + AB\mathbf{x} \\ &= \Theta(\mathbf{a} + A\mathbf{b}, AB)(\mathbf{x})\end{aligned}$$

donc

$$\Theta(\mathbf{a}, A) \circ \Theta(\mathbf{b}, B) = \Theta(\mathbf{a} + A\mathbf{b}, AB)$$

et alors

$$\Theta^{-1}[\Theta(\mathbf{a}, A) \circ \Theta(\mathbf{b}, B)] = (\mathbf{a} + A\mathbf{b}, AB).$$

(ii) (2.25) définit une loi interne sur $\mathbb{R}^n \times O(n)$. Vérifions qu'elle est associative. On a

$$\begin{aligned}(\mathbf{a}, A)((\mathbf{b}, B)(\mathbf{c}, C)) &= \Theta^{-1}[\Theta(\mathbf{a}, A) \circ \Theta((\mathbf{b}, B)(\mathbf{c}, C))] \\ &= \Theta^{-1}[\Theta(\mathbf{a}, A) \circ \Theta(\mathbf{b}, B) \circ \Theta(\mathbf{c}, C)] \\ &= \Theta^{-1}[\Theta((\mathbf{a}, A)(\mathbf{b}, B)) \circ \Theta(\mathbf{c}, C)] \\ &= ((\mathbf{a}, A)(\mathbf{b}, B))(\mathbf{c}, C)\end{aligned}$$

pour tous $(\mathbf{a}, A), (\mathbf{b}, B), (\mathbf{c}, C) \in \mathbb{R}^n \times O(n)$. La loi est associative. On peut aussi le vérifier en écrivant simplement

$$\begin{aligned}(\mathbf{a}, A)((\mathbf{b}, B)(\mathbf{c}, C)) &= (\mathbf{a}, A)(\mathbf{b} + B\mathbf{c}, BC) \\ &= (\mathbf{a} + A\mathbf{b} + AB\mathbf{c}, ABC) \\ &= (\mathbf{a} + A\mathbf{b}, AB)(\mathbf{c}, C) \\ &= ((\mathbf{a}, A)(\mathbf{b}, B))(\mathbf{c}, C).\end{aligned}$$

$(0, I) = \Theta^{-1}(Id_E)$ est élément neutre :

$$(0, I)(\mathbf{a}, A) = (\mathbf{a}, A) = (\mathbf{a}, A)(0, I) \quad \forall (\mathbf{a}, A) \in \mathbb{R}^n \times O(n).$$

Enfin $(-A^{-1}\mathbf{a}, A^{-1}) = \Theta^{-1}((\Theta(\mathbf{a}, A))^{-1})$ est l'inverse de $(\mathbf{a}, A)$ :

$$(-A^{-1}\mathbf{a}, A^{-1})(\mathbf{a}, A) = (0, I) = (\mathbf{a}, A)(-A^{-1}\mathbf{a}, A^{-1}).$$

Donc $\mathbb{R}^n \times O(n)$ est un groupe, et par définition de la loi (2.25) la bijection

$$\Theta : \mathbb{R}^n \times O(n) \to \mathcal{E}(E)$$

est un homorphisme de groupes, donc un isomorphisme. □

**Définition 2. 12.** *Le groupe $\mathbb{R}^n \times O(n)$ muni de la loi ci-dessus est appelé le groupe euclidien d'ordre $n$ et est noté $\mathcal{E}(n)$.*

# 2.4 Exponentielle d'une matrice.

## 2.4.1 Normes sur $M_n(\mathbb{C})$.

Soit $M_n(\mathbb{C})$ le $\mathbb{C}$-espace vectoriel des matrices $n\times n$ à coefficients complexes. Pour tout couple $(i,j)$, $1 \leqslant i,j \leqslant n$, soit $E_{i,j} \in M_n(\mathbb{C})$ la matrice dont le seul terme non nul est celui de la ligne $i$ et colonne $j$ qui vaut 1. Toute matrice $A = (a^i_j) \in M_n(\mathbb{C})$ s'écrivant de façon unique $A = \sum_{1\leqslant i,j\leqslant n} a^i_j\, E_{i,j}$,

$$(E_{1,1}, \cdots, E_{1,n}, E_{2,1}, \ldots, E_{2,n}, \ldots, E_{n,1}, \ldots E_{n,n})$$

est une base de $M_n(\mathbb{C})$, appelée *base canonique*. L'espace vectoriel $M_n(\mathbb{C})$ est donc de dimension $n^2$ et est isomorphe à $\mathbb{C}^{n^2}$ par l'application qui associe à une matrice $A = (a^i_j)$ le vecteur dont les composantes sont $a^1_1, \cdots, a^1_n, a^2_1, \ldots, a^2_n, \ldots, a^n_1, \ldots, a^n_n$.

On pose pour $A = (a^i_j) \in M_n(\mathbb{C})$ :

$$\|A\|_\infty = \sup_{i,j} |a^i_j|, \tag{2.26}$$

$$\|A\|_1 = \sum_{i,j} |a^i_j| \tag{2.27}$$

$$\|A\|_2 = \sqrt{\sum_{i,j} |a^i_j|^2}. \tag{2.28}$$

**Proposition 2. 10.** *Les formules (2.26), (2.27) et (2.28) définissent des normes sur l'espace vectoriel $M_n(\mathbb{C})$, et l'on a :*

$$\|A\|_\infty \leqslant \|A\|_2 \leqslant \|A\|_1 \leqslant n\|A\|_2 \leqslant n^2\|A\|_\infty \quad \forall\, A \in M_n(\mathbb{C}); \tag{2.29}$$

$$\|AB\|_1 \leqslant \|A\|_1\|B\|_1 \quad \forall\, A, B \in M_n(\mathbb{C}); \tag{2.30}$$

$$\|AB\|_2 \leqslant \|A\|_2\|B\|_2 \quad \forall\, A, B \in M_n(\mathbb{C}). \tag{2.31}$$

**Démonstration.**

Les applications $A \mapsto \|A\|_\infty$, $A \mapsto \|A\|_1$ et $A \mapsto \|A\|_2$ de $M_n(\mathbb{C})$ dans $[0, +\infty[$ vérifient

$$\|A\|_\infty = 0 \Leftrightarrow A = 0,$$

$$\|A\|_1 = 0 \Leftrightarrow A = 0,$$

$$\|A\|_2 = 0 \Leftrightarrow A = 0,$$

$$\|\lambda A\|_\infty = \sup_{i,j} |\lambda a^i_j| = |\lambda| \sup_{i,j} |a^i_j| = |\lambda|\, \|A\|_\infty,$$

$$\|\lambda A\|_1 = \sum_{i,j} |\lambda a^i_j| = |\lambda| \sum_{i,j} |a^i_j| = |\lambda|\, \|A\|_1,$$

$$\|\lambda A\|_2 = \sqrt{\sum_{i,j} |\lambda a^i_j|^2} = |\lambda| \sqrt{\sum_{i,j} |a^i_j|^2} = |\lambda|\, \|A\|_2$$

pour tous $\lambda \in \mathbb{C}, A = (a^i_j) \in M_n(\mathbb{C})$. Enfin

$$\|A + B\|_\infty = \sup_{i,j} |a^i_j + b^i_j| \leqslant \sup_{i,j} |a^i_j| + \sup_{i,j} |b^i_j| = \|A\|_\infty + \|B\|_\infty,$$

$$\|A+B\|_1=\sum_{i,j}|a_j^i+b_j^i|\leqslant\sum_{i,j}|a_j^i|+\sum_{i,j}|b_j^i|=\|A\|_1+\|B\|_1,$$

$$\begin{aligned}\|A+B\|_2^2 &= \sum_{i,j}|a_j^i+b_j^i|^2\\ &\leqslant \sum_{i,j}\left(|a_j^i|+|b_j^i|\right)^2\\ &\leqslant \sum_{i,j}|a_j^i|^2+\sum_{i,j}|b_j^i|^2+2\sum_{i,j}|a_j^i||b_j^i|\\ &\leqslant \sum_{i,j}|a_j^i|^2+\sum_{i,j}|b_j^i|^2+2\sqrt{\sum_{i,j}|a_j^i|^2}\sqrt{\sum_{i,j}|b_j^i|^2}\\ &\quad\text{(inégalité de Cauchy-Schwarz dans l'espace euclidien } \mathbb{R}^{n^2})\\ &\leqslant (\|A\|_2+\|B\|_2)^2\end{aligned}$$

*i.e.*

$$\|A+B\|_2\leqslant\|A\|_2+\|B\|_2$$

pour tous $A=(a_j^i), B=(b_j^i)\in M_n(\mathbb{C})$. Donc ce sont des normes.

Montrons maintenant (2.29).

- Il existe $i_0, j_0$ tels que $|a_{j_0}^{i_0}|=\sup_{i,j}|a_j^i|$, donc

$$\|A\|_\infty=\sup_{i,j}|a_j^i|\leqslant\sqrt{\sum_{i,j}|a_j^i|^2}=\|A\|_2.$$

- $\|A\|_2=\sqrt{\sum_{i,j}|a_j^i|^2}\leqslant\sqrt{\left(\sum_{i,j}|a_j^i|\right)^2}=\sum_{i,j}|a_j^i|=\|A\|_1.$
- $\|A\|_1=\sum_{i,j}|a_j^i|=\sum_{i,j}1\,|a_j^i|\leqslant\sqrt{\sum_{i,j}1^2}\sqrt{\sum_{i,j}|a_j^i|^2}=\sqrt{n^2}\|A\|_2=n\|A\|_2.$
- $\|A\|_2=\sqrt{\sum_{i,j}|a_j^i|^2}\leqslant\sqrt{n^2\|A\|_\infty^2}=n\|A\|_\infty$ donc $n\|A\|_2\leqslant n^2\|A\|_\infty$. Cela prouve (2.29).

Enfin, pour tous $A=(a_j^i), B=(b_j^i)\in M_n(\mathbb{C})$, on a en notant $C=AB=(c_j^i)$

$$\|C\|_1=\sum_{i,j}|c_j^i|=\sum_{i,j}\left|\sum_p a_p^ib_j^p\right|\leqslant\sum_{i,j,p}|a_p^i||b_j^p|\leqslant\sum_{i,j,p,q}|a_p^i||b_j^q|=\|A\|_1\|B\|_1$$

et (2.30) est prouvée. (2.31) se démontre de la même façon mais en utilisant l'inégalité de Cauchy-Schwarz :

$$\|C\|_2^2=\sum_{i,j}|c_j^i|^2=\sum_{i,j}\left|\sum_p a_p^ib_j^p\right|^2\leqslant\sum_{i,j}\left(\sum_p|a_p^i||b_j^p|\right)^2\leqslant\sum_{i,j}\left(\sum_p|a_p^i|^2\sum_q|b_j^q|^2\right)$$

$$=\sum_{i,j,p,q}|a_p^i|^2|b_j^q|^2=\left(\sum_{i,p}|a_p^i|^2\right)\left(\sum_{j,q}|b_j^q|^2\right)=\|A\|_2^2\|B\|_2^2.$$

□

**Remarque.** La norme $\|\cdot\|_\infty$ *ne vérifie pas* la condition analogue à (2.30) ou (2.31). Pour $A = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ et $B = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 0 \end{smallmatrix}\right)$, on a en effet $AB = \left(\begin{smallmatrix} 2 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, et

$$\|AB\|_\infty = 2 \ > \ \|A\|_\infty \|B\|_\infty = 1 \cdot 1.$$

**Corollaire 1.** *Soit $(A_k)_{k\in\mathbb{N}}$, $A_k = (a^i_j(k)) \in M_n(\mathbb{C})$, une suite de matrices et $B = (b^i_j) \in M_n(\mathbb{C})$. Les propriétés suivantes sont équivalentes.*

*(i) La suite $(A_k)_{k\in\mathbb{N}}$ converge vers $B$ quand $k \to +\infty$ dans l'espace normé $(M_n(\mathbb{C}), \|\cdot\|_1)$.*

*(ii) La suite $(A_k)_{k\in\mathbb{N}}$ converge vers $B$ quand $k \to +\infty$ dans l'espace normé $(M_n(\mathbb{C}), \|\cdot\|_2)$.*

*(iii) La suite $(A_k)_{k\in\mathbb{N}}$ converge vers $B$ quand $k \to +\infty$ dans l'espace normé $(M_n(\mathbb{C}), \|\cdot\|_\infty)$.*

*(iv) Pour tous $i, j$ le coefficient $a^i_j(k)$ tend vers $b^i_j$ quand $k \to +\infty$.*

**Démonstration.**

(i) $\Rightarrow$ (ii) :

$$\|A_k - B\|_2 \leqslant \|A_k - B\|_1 \to 0 \text{ quand } k \to +\infty.$$

(ii) $\Rightarrow$ (iii) :

$$\|A_k - B\|_\infty \leqslant \|A_k - B\|_2 \to 0 \text{ quand } k \to +\infty.$$

(iii) $\Rightarrow$ (iv) :

$$\left|a^i_j(k) - b^i_j\right| \leqslant \|A_k - B\|_\infty \to 0 \text{ quand } k \to +\infty.$$

(iv) $\Rightarrow$ (i) :

$$\|A_k - B\|_1 = \sum_{i,j} \left|a^i_j(k) - b^i_j\right| \to 0 \text{ quand } k \to +\infty.$$

□

**Notation.** Si les conditions équivalentes du corollaire précédent sont vérifiées, on notera simplement $A_k \to B$ ou $B = \lim A_k$.

Dans toute la suite, on munit $M_n(\mathbb{C})$ de la norme $\|\cdot\|_1$ en raison de l'inégalité (2.30). On aurait aussi bien pu choisir la norme $\|\cdot\|_2$ puisqu'elle satisfait (2.31). Les résultats seraient les mêmes d'après le Corollaire 1.

**Corollaire 2.** *On munit $M_n(\mathbb{C})$ de la norme $\|\cdot\|_1$. L'application bilinéaire $(A, B) \mapsto AB$ de $M_n(\mathbb{C}) \times M_n(\mathbb{C})$ dans $M_n(\mathbb{C})$ est continue, i.e. la multiplication matricielle est continue.*

**Démonstration.**

Il s'agit de montrer que si $A_k \to A$ et $B_k \to B$, alors $C_k = A_k B_k \to C = AB$. Si $A_k = (a^i_j(k)), A = (a^i_j), B_k = (b^i_j(k)), B = (b^i_j), C_k = (c^i_j(k)), C = (c^i_j)$, on a $c^i_j(k) = \sum_{p=1}^n a^i_p(k) b^p_j(k)$ et $c^i_j = \sum_{p=1}^n a^i_p b^p_j$. Or pour tous $i, j$ $a^i_p(k) \to a^i_p$ et $b^p_j(k) \to b^p_j$, donc $c^i_j(k) \to c^i_j$. Cela étant vrai pour tous $i, j$, on a bien $C_k \to C$. □

**Corollaire 3.** *Soit $(A_k)_{k\in\mathbb{N}}$, $A_k = (a^i_j(k))$ une suite de matrices de $M_n(\mathbb{C})$. Si la série $\sum_{k=0}^{+\infty} \|A_k\|_1$ converge, i.e. la série $\sum_{k=0}^{+\infty} A_k$ est absolument convergente dans*

*l'espace normé* $(M_n(\mathbb{C}), \|\cdot\|_1)$, *alors la série* $\sum_{k=0}^{+\infty} A_k$ *converge dans l'espace normé* $(M_n(\mathbb{C}), \|\cdot\|_1)$. *On a de plus*

$$\left\|\sum_{k=0}^{+\infty} A_k\right\|_1 \leqslant \sum_{k=0}^{+\infty} \|A_k\|_1 . \tag{2.32}$$

**Démonstration.**

Pour tous $i, j$ fixés, on a $|a_j^i(k)| \leqslant \|A_k\|_1 \ \forall k \in \mathbb{N}$. La série numérique $\sum_{k=0}^{+\infty} a_j^i(k)$ est donc absolument convergente, et on sait que cela implique sa convergence. Posons $b_j^i = \sum_{k=0}^{+\infty} a_j^i(k)$, et soit $B$ la matrice $(b_j^i)$. Alors

$$\left\|\sum_{k=0}^{N} A_k - B\right\|_1 \to 0 \text{ quand } N \to +\infty$$

d'après le Corollaire 1 puisque pour tous $i, j$ fixés, $\sum_{k=0}^{N} a_j^i(k) \to b_j^i$ quand $N \to +\infty$. Cela prouve que la série $\sum_{k=0}^{+\infty} A_k$ converge dans l'espace normé $(M_n(\mathbb{C}), \|\cdot\|_1)$ et que sa somme $\sum_{k=0}^{+\infty} A_k$ est $B$.

D'autre part on a pour tout $N$,

$$\left|\left\|\sum_{k=0}^{N} A_k\right\|_1 - \|B\|_1\right| \leqslant \left\|\sum_{k=0}^{N} A_k - B\right\|_1 ,$$

donc

$$\left\|\sum_{k=0}^{N} A_k\right\|_1 \to \|B\|_1 \text{ quand } N \to +\infty.$$

Or

$$\left\|\sum_{k=0}^{N} A_k\right\|_1 \leqslant \sum_{k=0}^{N} \|A_k\|_1 \leqslant \sum_{k=0}^{+\infty} \|A_k\|_1 \ \forall N \in \mathbb{N},$$

donc par passage à la limite on obtient (2.32). □

### 2.4.2 Exponentielle sur $M_n(\mathbb{C})$.

**Proposition 2. 11.** *Soit* $A \in M_n(\mathbb{C})$. *La série* $\sum_{k=0}^{+\infty} \frac{1}{k!} A^k$ *converge dans l'espace normé* $(M_n(\mathbb{C}), \|\cdot\|_1)$.

**Démonstration.**

Il suffit d'après le Corollaire 3 de la Proposition 2.10 de montrer que la série est absolument convergente. Or d'après (2.30),

$$\|A^k\|_1 \leqslant \|A\|_1^k \quad \forall k \in \mathbb{N}.$$

La série $\sum_{k=0}^{+\infty} \frac{1}{k!} \|A\|_1^k$ étant convergente puisque la série exponentielle $\sum_{k=0}^{+\infty} \frac{1}{k!} z^k$ a pour rayon de convergence $+\infty$, le résultat en découle. □

**Définition 2. 13.** *Soit* $A \in M_n(\mathbb{C})$. *La somme de la série* $\sum_{k=0}^{+\infty} \frac{1}{k!} A^k$ *est appelée exponentielle de* $A$ *et notée* $e^A$.

### 2.4.3 Propriétés de l'exponentielle d'une matrice.

**Proposition 2. 12.**

$$e^{PAP^{-1}} = Pe^A P^{-1} \quad \forall A \in M_n(\mathbb{C}), \quad \forall P \in GL(n, \mathbb{C}).$$

**Démonstration.**

Pour tout $N \in \mathbb{N}$,

$$\sum_{k=0}^{N} \frac{1}{k!}(PAP^{-1})^k = \sum_{k=0}^{N} \frac{1}{k!} PA^k P^{-1} = P\left(\sum_{k=0}^{N} \frac{1}{k!} A^k\right) P^{-1}.$$

La multiplication matricielle étant continue, on obtient $e^{PAP^{-1}} = Pe^A P^{-1}$ par passage à la limite quand $N \to +\infty$. □

**Proposition 2. 13.** *Soient $A, B \in M_n(\mathbb{C})$. Si $A$ et $B$ commutent, i.e. $AB = BA$, alors*

$$e^{A+B} = e^A e^B.$$

**Démonstration.**

Ecrivons pour $N \geqslant 1$ :

$$\left(\sum_{p=0}^{N} \frac{1}{p!} A^p\right)\left(\sum_{q=0}^{N} \frac{1}{q!} A^q\right) = \Sigma_N + R_N \tag{2.33}$$

avec

$$\Sigma_N = \sum_{p+q \leqslant N} \frac{1}{p!}\frac{1}{q!} A^p B^q$$

$$R_N = \sum_{\substack{N+1 \leqslant p+q \leqslant 2N \\ 1 \leqslant p,q \leqslant N}} \frac{1}{p!}\frac{1}{q!} A^p B^q$$

On a en posant $M = \sup(\|A\|_1, \|B\|_1)$ :

$$\begin{aligned}
\|R_N\|_1 &\leqslant \sum_{\substack{N+1 \leqslant p+q \leqslant 2N \\ 1 \leqslant p,q \leqslant N}} \frac{1}{p!}\frac{1}{q!} \|A^p\|_1 \|B^q\|_1 \\
&\leqslant \sum_{\substack{N+1 \leqslant p+q \leqslant 2N \\ 1 \leqslant p,q \leqslant N}} \frac{1}{p!}\frac{1}{q!} \|A\|_1^p \|B\|_1^q \\
&\leqslant \sum_{\substack{N+1 \leqslant p+q \leqslant 2N \\ 1 \leqslant p,q \leqslant N}} \frac{1}{p!}\frac{1}{q!} M^{p+q} = \sum_{k=N+1}^{2N} M^k \left(\sum_{\substack{p+q=k \\ 1 \leqslant p,q \leqslant N}} \frac{1}{p!}\frac{1}{q!}\right) \\
&\leqslant \sum_{k=N+1}^{2N} M^k \left(\sum_{p+q=k} \frac{1}{p!}\frac{1}{q!}\right) = \sum_{k=N+1}^{2N} M^k \frac{2^k}{k!} \\
&\leqslant \sum_{k=N+1}^{+\infty} M^k \frac{2^k}{k!} \to 0 \text{ quand } N \to +\infty
\end{aligned}$$

puisque $\sum_{k=N+1}^{+\infty} M^k \frac{2^k}{k!}$ est le reste d'indice $N$ de la série convergente $\sum_{k=0}^{+\infty} M^k \frac{2^k}{k!}$.

D'autre part,

$$\begin{aligned}\Sigma_N &= \sum_{k=0}^{N}\left(\sum_{p+q=k}\frac{1}{p!}\frac{1}{q!}A^pB^q\right)\\ &= \sum_{k=0}^{N}\frac{1}{k!}(A+B)^k\end{aligned}$$

puisque la formule du binôme s'applique à $(A+B)^k$ du fait que $A$ et $B$ commutent. On a donc

$$\Sigma_N \to e^{A+B} \text{ quand } N \to +\infty.$$

Le second membre de la formule (2.33) tend donc vers $e^{A+B}$ quand $N \to +\infty$.

Le premier membre tend vers $e^A e^B$ puisque la multiplication matricielle est continue pour la norme $\|\cdot\|_1$. Par passage à la limite dans la formule (2.33) on obtient donc $e^{A+B} = e^A e^B$. □

**Corollaire.** *Pour tout $A \in M_n(\mathbb{C})$, la matrice $e^A$ est inversible et $(e^A)^{-1} = e^{-A}$.*

**Démonstration.**

On a en effet $e^A e^{-A} = e^{-A}e^A = e^0 = I$. □

**Définition 2. 14.** *Une application $t \mapsto A(t) = (a_j^i(t))$ de $\mathbb{R}$ dans $M_n(\mathbb{C})$ est dite dérivable au point $t_0 \in \mathbb{R}$ si chaque fonction $a_j^i$ l'est. On note alors $A'(t_0) = ((a_j^i)'(t_0))$ ou encore $\left(\frac{d}{dt}A\right)_{t=t_0} = \left(\left(\frac{d}{dt}a_j^i\right)_{t=t_0}\right)$.*

**Proposition 2. 14.** *Une application $t \mapsto A(t) = (a_j^i(t))$ de $\mathbb{R}$ dans $M_n(\mathbb{C})$ est dérivable au point $t_0 \in \mathbb{R}$ si et seulement si le rapport*

$$\frac{A(t_0+s)-A(t_0)}{s}$$

*possède une limite $B$ dans l'espace normé $(M_n(\mathbb{C}), \|\cdot\|_1)$ quand $s \to 0$. On a alors $B = A'(t_0)$.*

**Démonstration.**

On a :

$$\frac{A(t_0+s)-A(t_0)}{s} = \left(\frac{a_j^i(t_0+s)-a_j^i(t_0)}{s}\right).$$

Dire que la matrice $\frac{A(t_0+s)-A(t_0)}{s}$ possède une limite $B = (b_j^i)$ dans l'espace normé $(M_n(\mathbb{C}), \|\cdot\|_1)$ signifie donc que

$$\left\|\left(\frac{a_j^i(t_0+s)-a_j^i(t_0)}{s} - b_j^i\right)\right\|_1 \to 0 \text{ quand } s \to 0$$

et ceci équivaut à dire que pour tous $i, j$

$$\frac{a_j^i(t_0+s)-a_j^i(t_0)}{s} - b_j^i \to 0 \text{ quand } s \to 0$$

*i.e.* $a^i_j$ est dérivable au point $t_0$, de dérivée $(a^i_j)'(t_0) = b^i_j$. □

**Théorème 2. 11.** *Soit $A \in M_n(\mathbb{C})$. L'application $t \mapsto e^{tA}$ de $\mathbb{R}$ dans $M_n(\mathbb{C})$ est dérivable sur $\mathbb{R}$ et l'on a pour tout $t \in \mathbb{R}$*

$$\frac{d}{dt}e^{tA} = Ae^{tA} = e^{tA}A. \tag{2.34}$$

*Elle vérifie*

$$e^{(t+s)A} = e^{tA}e^{sA} \quad \forall\, t, s \in \mathbb{R}. \tag{2.35}$$

**Démonstration.**

Comme
$$\frac{e^{(t+s)A} - e^{tA}}{s} = e^{tA}\,\frac{e^{sA} - I}{s} = \frac{e^{sA} - I}{s}\,e^{tA},$$
il suffit de démontrer que l'application $t \mapsto e^{tA}$ est dérivable au point $t = 0$ et que
$$\left(\frac{d}{dt}e^{tA}\right)_{t=0} = A.$$

Considérons donc $\frac{e^{sA}-I}{s}$. On a
$$\frac{e^{sA} - I}{s} = A + \frac{s}{2!}A^2 + \frac{s^2}{3!}A^3 + \cdots = A + \sum_{k=2}^{+\infty}\frac{s^{k-1}}{k!}A^k,$$
donc
$$\frac{e^{sA} - I}{s} - A = \sum_{k=2}^{+\infty}\frac{s^{k-1}}{k!}A^k.$$

On a alors :
$$\begin{aligned}
\left\|\frac{e^{sA} - I}{s} - A\right\|_1 &= \left\|\sum_{k=2}^{+\infty}\frac{s^{k-1}}{k!}A^k\right\|_1 \\
&\leqslant \sum_{k=2}^{+\infty}\frac{|s|^{k-1}}{k!}\|A^k\|_1 \\
&\leqslant \sum_{k=2}^{+\infty}\frac{|s|^{k-1}}{k!}\|A\|_1^k = |s|\,\|A\|_1^2\sum_{k=2}^{+\infty}\frac{|s|^{k-2}}{k!}\|A\|_1^{k-2} \\
&\leqslant |s|\,\|A\|_1^2\sum_{k=2}^{+\infty}\frac{|s|^{k-2}}{(k-2)!}\|A\|_1^{k-2} = |s|\,\|A\|_1^2\,e^{|s|\,\|A\|_1}.
\end{aligned}$$

Or
$$|s|\,\|A\|_1^2\,e^{|s|\,\|A\|_1} \to 0 \text{ quand } s \to 0.$$

Donc
$$\left\|\frac{e^{sA} - I}{s} - A\right\|_1 \to 0 \text{ quand } s \to 0$$
et l'application $t \mapsto e^{tA}$ est dérivable au point $t = 0$ avec
$$\left(\frac{d}{dt}e^{tA}\right)_{t=0} = A.$$

□

## 2.5 Exercices.

**Exercice 2.1.**

Trouver les idéaux bilatères de l'anneau $M_n(\mathbb{C})$.

**Indication.**

Soit $\mathcal{I}$ un idéal bilatère de $M_n(\mathbb{C})$. Supposons $\mathcal{I} \neq \{0\}$. Soit $A \in \mathcal{I}, A \neq 0$. Si $A$ est inversible, $I = AA^{-1} \in \mathcal{I}$, d'où $\mathcal{I} = M_n(\mathbb{C})$. Si $A$ n'est pas inversible, le rang $r$ de $A$ est $< n$, donc $1 \leqslant r < n$. On sait que toute matrice de rang $r$ est équivalente à la matrice $J_r$ (2.8). Il existe donc des matrices inversibles $P, Q$ telles que $J_r = QAP$. Comme $A \in \mathcal{I}$ et que $\mathcal{I}$ est un idéal bilatère, on a alors $J_r \in \mathcal{I}$. Maintenant, pour $1 \leqslant i \leqslant n$, soit $E_{i,i}$ la matrice dont le seul terme non nul est celui de la ligne $i$ et colonne $i$ qui vaut 1. Il est immédiat que $E_{1,1} = E_{1,1}J_r$. Donc $E_{1,1} \in \mathcal{I}$. Pour $i$ quelconque, $E_{i,i}$ est de rang 1 comme $E_{1,1}$. Elles sont donc toutes deux équivalentes à $J_1$, et cela implique en particulier que $E_{i,i}$ est équivalente à $E_{1,1}$. Il existe donc $R, S$ inversibles telles que $E_{i,i} = RE_{1,1}S$. Comme $E_{1,1} \in \mathcal{I}$, on a alors $E_{i,i} \in \mathcal{I}$. Cela étant vrai pour tout $i$, la matrice $\sum_{i=1}^{n} E_{i,i}$ appartient à $\mathcal{I}$. Mais cette matrice n'est autre que la matrice identité $I$. Il en résulte alors que $\mathcal{I} = M_n(\mathbb{C})$. En conclusion, les seuls idéaux bilatères de $M_n(\mathbb{C})$ sont $\{0\}$ et $M_n(\mathbb{C})$.

**Exercice 2.2.**

Montrer que l'endomorphisme de $\mathbb{R}^2$ dont la matrice dans la base orthonormée canonique est $\left(\begin{smallmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{smallmatrix}\right)$, $(\theta \in \mathbb{R})$ est la symétrie orthogonale par rapport à la droite de pente $\tan\frac{\theta}{2}$ dont un vecteur directeur est $\mathbf{v} = \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} \end{pmatrix}$.

**Exercice 2.3.**

Soit $f$ l'endomorphisme de $\mathbb{R}^3$ dont la matrice dans la base orthonormée canonique est

$$A = \frac{1}{3}\begin{pmatrix} 1 & 1-\sqrt{3} & 1+\sqrt{3} \\ 1+\sqrt{3} & 1 & 1-\sqrt{3} \\ 1-\sqrt{3} & 1+\sqrt{3} & 1 \end{pmatrix}$$

(i) Montrer que $f$ est une rotation.

(ii) Calculer un vecteur directeur de l'axe de rotation et l'angle de rotation.

(iii) Calculer une matrice antisymétrique $X$ telle que l'on ait $A = e^{\theta X}$ pour un $\theta \in \mathbb{R}$ que l'on précisera.

**Indication.**

(i) On a facilement ${}^tA\,A = I$ donc $A$ est orthogonale. Pour calculer $\det A$, on remplace d'abord la 1ère colonne par la somme de 3 colonnes. *i.e.* on fait $c'_1 = c_1 + c_2 + c_3$.

$$\det A = \frac{1}{27}\begin{vmatrix} 1 & 1-\sqrt{3} & 1+\sqrt{3} \\ 1+\sqrt{3} & 1 & 1-\sqrt{3} \\ 1-\sqrt{3} & 1+\sqrt{3} & 1 \end{vmatrix} = \frac{1}{9}\begin{vmatrix} 1 & 1-\sqrt{3} & 1+\sqrt{3} \\ 1 & 1 & 1-\sqrt{3} \\ 1 & 1+\sqrt{3} & 1 \end{vmatrix}$$

$$= \frac{1}{9}\begin{vmatrix} 1 & 1-\sqrt{3} & 1+\sqrt{3} \\ 0 & \sqrt{3} & -2\sqrt{3} \\ 0 & 2\sqrt{3} & -\sqrt{3} \end{vmatrix} = 1\,.$$

D'où $A \in SO(3)$.

(ii) L'axe est l'ensemble des vecteurs invariants par $A$. Si $\mathbf{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, on a

$$\begin{aligned}
A\mathbf{v} = \mathbf{v} \quad &\Leftrightarrow \quad \begin{cases} -2x+(1-\sqrt{3})y+(1+\sqrt{3})z=0 & (L_1) \\ (1+\sqrt{3})x-2y+(1-\sqrt{3})z=0 & (L_2) \\ (1-\sqrt{3})x+(1+\sqrt{3})y-2z=0 & (L_3) \end{cases} \\
&\Leftrightarrow \quad \begin{cases} -2x+(1-\sqrt{3})y+(1+\sqrt{3})z=0 & (L_1) \\ -6y+6z=0 & (L_2'=2L_2+(1+\sqrt{3})L_1) \\ 6y-6z=0 & (L_3'=2L_3+(1-\sqrt{3})L_1) \end{cases} \\
&\Leftrightarrow \quad x = y = z \\
&\Leftrightarrow \quad \mathbf{v} = x \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.
\end{aligned}$$

On peut prendre comme vecteur directeur de l'axe le vecteur normé $\mathbf{k} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Soit $\theta$ la détermination principale de l'angle ($-\pi < \theta \leqslant \pi$). On a $1 + \cos\theta = \operatorname{Tr} A = 1$, donc $\cos\theta = 0$, *i.e.* $\theta = \pm\frac{\pi}{2}$. On sait que le signe de $\theta$ est celui du produit mixte $[\mathbf{v}, A\mathbf{v}, \mathbf{k}]$ où $\mathbf{v}$ est un vecteur quelconque non colinéaire à $\mathbf{k}$. En prenant $\mathbf{v} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, on obtient $\theta = \frac{\pi}{2}$ puisque

$$[\mathbf{v}, A\mathbf{v}, \mathbf{k}] = \frac{1}{3\sqrt{3}} \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1+\sqrt{3} & 1 \\ 0 & 1-\sqrt{3} & 1 \end{vmatrix} = \frac{2}{3}.$$

(iii) Construisons une base orthonormée directe $(\mathbf{f_1}, \mathbf{f_2}, \mathbf{f_3})$ telle que $\mathbf{f_3} = \mathbf{k}$. On peut prendre par exemple $\mathbf{f_1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ et $\mathbf{f_2} = \mathbf{f_3} \wedge \mathbf{f_1} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$. Alors, avec $\theta = \frac{\pi}{2}$,

$$B = \mathcal{M}(f, (\mathbf{f_1}, \mathbf{f_2}, \mathbf{f_3})) = P^{-1}AP = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

où

$$P = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{3} & 1 & \sqrt{2} \\ -\sqrt{3} & 1 & \sqrt{2} \\ 0 & -2 & \sqrt{2} \end{pmatrix}.$$

On sait que $P$ est orthogonale puisque c'est une matrice de changement de bases orthonormées, donc $P^{-1} = {}^tP$. D'autre part, on a si

$$J = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$e^{tJ} = \begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \forall t \in \mathbb{R}.$$

Donc $B = e^{\theta J}$. Alors $A = Pe^{\theta J}P^{-1} = e^{\theta PJP^{-1}} = e^{\theta X}$ avec

$$X = PJ\,{}^tP = \frac{1}{\sqrt{3}} \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}.$$

**Exercice 2.4.**

Soit $A = (a^i_j) \in SO(3)$. Montrer que $\left|\sum_{i,j=1}^3 a^i_j\right| \leqslant 3$.

**Indication.**

Utiliser l'inégalité de Cauchy-Schwarz et le vecteur $\mathbf{u} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

**Exercice 2.5.**

Soit $E$ un espace euclidien et $f$ un endomorphisme $\neq 0$ de $E$ qui *conserve l'orthogonalité*, *i.e.* ayant la propriété suivante :

$$\forall \mathbf{x}, \mathbf{y} \in E, \quad (\mathbf{x}|\mathbf{y}) = 0 \Rightarrow (f(\mathbf{x})|f(\mathbf{y})) = 0. \tag{2.36}$$

Montrer qu'il existe un réel $\lambda > 0$ tel que

$$(f(\mathbf{x})|f(\mathbf{y})) = \lambda^2 (\mathbf{x}|\mathbf{y}) \quad \forall \mathbf{x}, \mathbf{y} \in E.$$

Quelle est l'interprétation géométrique de $f$ ?

**Indication.**

Soit $\mathcal{B} = (\mathbf{e_1}, \dots, \mathbf{e_n})$ une base orthonormée de $E$. D'après l'hypothèse (2.36)

$$(f(\mathbf{e_i})|f(\mathbf{e_j})) = 0 \quad \forall i, j \quad 1 \leqslant i, j \leqslant n.$$

Montrons que

$$\|f(\mathbf{e_i})\| = \|f(\mathbf{e_j})\| \quad \forall i, j \quad 1 \leqslant i, j \leqslant n.$$

On a

$$\begin{aligned} \|f(\mathbf{e_i})\|^2 - \|f(\mathbf{e_j})\|^2 &= (f(\mathbf{e_i})|f(\mathbf{e_i})) - (f(\mathbf{e_j})|f(\mathbf{e_j})) \\ &= (f(\mathbf{e_i}) - f(\mathbf{e_j})|f(\mathbf{e_i}) + f(\mathbf{e_j})) \\ &= (f(\mathbf{e_i} - \mathbf{e_j})|f(\mathbf{e_i} + \mathbf{e_j})) \\ &= 0 \end{aligned}$$

d'après l'hypothèse (2.36) puisque

$$(\mathbf{e_i} - \mathbf{e_j}|\mathbf{e_i} + \mathbf{e_j}) = \|\mathbf{e_i}\|^2 - \|\mathbf{e_j}\|^2 = 0.$$

Si $\lambda$ désigne $\|f(\mathbf{e_1})\|$, on a donc $\|f(\mathbf{e_i})\| = \lambda \ \forall i,\ 1 \leqslant i \leqslant n$. En particulier, $\lambda > 0$ puisque $f \neq 0$. Pour tous $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e_i} \in E, \quad \mathbf{y} = \sum_{i=1}^n y_i \mathbf{e_i} \in E$ on a alors

$$(f(\mathbf{x})|f(\mathbf{y})) = \sum_{i,j} x_i y_j (f(\mathbf{e_i})|f(\mathbf{e_j})) = \sum_i x_i y_i \|f(\mathbf{e_i})\|^2 = \lambda^2 (\mathbf{x}|\mathbf{y}).$$

Si l'on introduit $g = \frac{1}{\lambda} f$, alors $(g(\mathbf{x})|g(\mathbf{y})) = (\mathbf{x}|\mathbf{y}) \quad \forall \mathbf{x}, \mathbf{y} \in E$, donc $g$ est un endomorphisme isométrique de $E$. L'endomorphisme $f = \lambda g$ est une *similitude* vectorielle de $E$.

**Exercice 2.6.**

Soit $G$ le sous-ensemble de $GL(n+1, \mathbb{R})$ défini par :

$$G = \left\{ \begin{pmatrix} & & & a_1 \\ & A & & \vdots \\ & & & a_n \\ 0 & \dots & 0 & 1 \end{pmatrix} ; \quad A \in O(n), \quad a_1, \dots, a_n \in \mathbb{R} \right\}.$$

Montrer que $G$ est un sous-groupe de $GL(n+1,\mathbb{R})$ isomorphe au groupe euclidien $\mathcal{E}(n)$.

**Exercice 2.7.**

Soit $E$ l'espace vectoriel des polynômes de degré $\leqslant n$ et à coefficients réels.

(i) On pose pour $P,Q \in E$ : $(P|Q) = \int_{-1}^{1} P(x)Q(x)dx$. Montrer que cette formule définit un produit scalaire sur $E$.

(ii) Pour $0 \leqslant k \leqslant n$, on pose

$$P_k(x) = \frac{1}{2^k k!}\frac{d^k}{dx^k}(x^2-1)^k.$$

Montrer que $(\sqrt{\frac{2k+1}{2}}\,P_k;\ 0 \leqslant k \leqslant n)$ est une base orthonormée de l'espace euclidien $E$ (pour le produit scalaire de (i)). Les $P_k$ sont les polynômes de Legendre.

**Indication.**

(ii) On vérifie facilement que $P_k$ est un polynôme de degré $k$. Pour montrer que $\{P_k\ ;\ 0 \leqslant k \leqslant n\}$ est un système orthogonal de $E$, il suffit donc de vérifier que pour tout $k$ on a

$$(P_k|x^\ell) = 0 \quad \forall \ell \quad 0 \leqslant \ell < k.$$

Or 1 et $-1$ sont des zéros d'ordre $k$ du polynôme $Q(x) = (x^2-1)^k$, donc des zéros d'ordre $k-p$ du polynôme $Q^{(p)}(x) = \frac{d^p}{dx^p}(x^2-1)^k$ pour $0 \leqslant p < k$. Par intégrations par parties successives on obtient alors pour $0 \leqslant \ell \leqslant k$ :

$$\begin{aligned}
(P_k|x^\ell) &= \frac{1}{2^k k!}\int_{-1}^{1} x^\ell \frac{d^k}{dx^k}(x^2-1)^k\,dx \\
&= -\frac{1}{2^k k!}\ell\int_{-1}^{1} x^{\ell-1}\frac{d^{k-1}}{dx^{k-1}}(x^2-1)^k\,dx \\
&= \frac{1}{2^k k!}\ell(\ell-1)\int_{-1}^{1} x^{\ell-2}\frac{d^{k-2}}{dx^{k-2}}(x^2-1)^k\,dx \\
&= \cdots \\
&= \frac{1}{2^k k!}(-1)^\ell \ell!\int_{-1}^{1}\frac{d^{k-\ell}}{dx^{k-\ell}}(x^2-1)^k\,dx\,.
\end{aligned}$$

• Pour $\ell < k$,

$$\int_{-1}^{1}\frac{d^{k-\ell}}{dx^{k-\ell}}(x^2-1)^k\,dx = \left[\frac{d^{k-\ell-1}}{dx^{k-\ell-1}}(x^2-1)^k\right]_{-1}^{1} = 0,$$

donc $(P_k|x^\ell) = 0 \quad \forall \ell \ \ 0 \leqslant \ell < k$ et $\{P_k\ ;\ 0 \leqslant k \leqslant n\}$ est un système orthogonal de $E$.

• Pour $\ell = k$,

$$\begin{aligned}
\int_{-1}^{1}(x^2-1)^k\,dx &= 2(-1)^k\int_{0}^{1}(1-x^2)^k\,dx \\
&= 2(-1)^k\int_{0}^{\frac{\pi}{2}}\cos^{2k+1}t\,dt \quad \text{en posant } x=\sin t \\
&= 2(-1)^k I_{2k+1}
\end{aligned}$$

où $I_{2k+1}$ désigne l'intégrale de Wallis $\int_0^{\frac{\pi}{2}}\cos^{2k+1}t\ dt$. Un calcul classique (où l'on montre d'abord par intégration par parties de $I_{2k+1}$ en posant $u = \cos^{2k}t$ et

$dv = \cos t\, dt$ la relation de récurrence $(2k+1)I_{2k+1} = 2kI_{2k-1}$ $(k \geqslant 1)$) montre que

$$I_{2k+1} = \frac{2^{2k}(k!)^2}{(2k+1)!}.$$

Donc

$$(P_k|x^k) = 2^{k+1}\,\frac{(k!)^2}{(2k+1)!}.$$

Or le terme de degré $k$ de $P_k$ est

$$\frac{1}{2^k k!}2k\,(2k-1)\cdots(k+1)\,x^k = \frac{(2k)!}{2^k(k!)^2}\,x^k.$$

Donc, on obtient compte tenu de $(P_k|x^\ell) = 0 \quad \forall \ell \;\; 0 \leqslant \ell < k$ :

$$\|P_k\|^2 = (P_k|P_k) = (P_k \mid \frac{(2k)!}{2^k(k!)^2}\,x^k\,) = \frac{2}{(2k+1)}.$$

Ainsi $(\sqrt{\frac{2k+1}{2}}\,P_k;\; 0 \leqslant k \leqslant n)$ est une base orthonormée de $E$.

**Exercice 2.8.**

Soit

$$A = \begin{pmatrix} -1 & 1 & 1 \\ -5 & 21 & 17 \\ 6 & -26 & -21 \end{pmatrix}.$$

(i) Trouver les valeurs propres de $A$. $A$ est-elle diagonalisable?

(ii) Trouver une base $\{\mathbf{b}, \mathbf{c}\}$ de $\operatorname{Ker} A^2$ telle que $\mathbf{b} = A\mathbf{c}$. En déduire une matrice $P$ telle que

$$P^{-1}AP = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

(iii) Calculer $e^{tA}$ pour $t \in \mathbb{R}$.

**Indication.**

(i) Le polynôme caractéristisque de $A$ est

$$P_A(\lambda) = \begin{vmatrix} -1-\lambda & 1 & 1 \\ -5 & 21-\lambda & 17 \\ 6 & -26 & -21-\lambda \end{vmatrix}.$$

Pour le calculer, on remplace d'abord la 1ère colonne par la 1ère colonne plus la 2ème moins la 3ème, *i.e.* on fait $c_1' = c_1 + c_2 - c_3$.

$$\begin{aligned} P_A(\lambda) &= \begin{vmatrix} -1-\lambda & 1 & 1 \\ -1-\lambda & 21-\lambda & 17 \\ 1+\lambda & -26 & -21-\lambda \end{vmatrix} \\ &= -(1+\lambda)\begin{vmatrix} 1 & 1 & 1 \\ 1 & 21-\lambda & 17 \\ -1 & -26 & -21-\lambda \end{vmatrix} \\ &= -(1+\lambda)\begin{vmatrix} 1 & 0 & 0 \\ 1 & 20-\lambda & 16 \\ -1 & -25 & -20-\lambda \end{vmatrix} \\ &= -(1+\lambda)\begin{vmatrix} 20-\lambda & 16 \\ -25 & -20-\lambda \end{vmatrix} \\ &= -(1+\lambda)\lambda^2. \end{aligned}$$

Les valeurs propres sont donc $-1$ (simple) et $0$ (double). Le sous-espace propre associé à la valeur propre 0 est $\operatorname{Ker} A = \mathbb{R}\begin{pmatrix} 1 \\ -3 \\ 4 \end{pmatrix}$. Il est de dimension 1, donc $A$ n'est pas diagonalisable.

(ii) $A^2 = \begin{pmatrix} 2 & -6 & -5 \\ 2 & -6 & -5 \\ -2 & 6 & 5 \end{pmatrix}$, donc

$$\mathbf{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \operatorname{Ker} A^2 \Leftrightarrow 2x - 6y - 5z = 0 \Leftrightarrow \mathbf{v} = x\begin{pmatrix} 1 \\ 0 \\ \frac{2}{5} \end{pmatrix} + y\begin{pmatrix} 0 \\ 1 \\ -\frac{6}{5} \end{pmatrix}.$$

$\left(\begin{pmatrix} 5 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 5 \\ -6 \end{pmatrix}\right)$ est une base de Ker $A^2$. Posons $\mathbf{c} = \begin{pmatrix} 5 \\ 0 \\ 2 \end{pmatrix}$. On a $A\mathbf{c} = \begin{pmatrix} -3 \\ 9 \\ -12 \end{pmatrix}$. On pose $\mathbf{b} = A\mathbf{c}$. Comme $\mathbf{b}, \mathbf{c}$ sont indépendants et que $\dim \operatorname{Ker} A^2 = 2$, ils forment une base de $\operatorname{Ker} A^2$ répondant à la question. On a $A\mathbf{b} = A^2\mathbf{c} = 0$ puisque $\mathbf{c} \in \operatorname{Ker} A^2$, *i.e.* $\mathbf{b}$ est vecteur propre de $A$ pour la valeur propre 0. D'autre part on calcule facilement le sous-espace propre pour la valeur propre $-1$ : $\operatorname{Ker}(A+1) = \mathbb{R}\mathbf{a}$ avec $\mathbf{a} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$. Alors $\mathcal{B}' = (\mathbf{a}, \mathbf{b}, \mathbf{c})$ est une base de $\mathbb{R}^3$ et la matrice de passage $P$ de la base canonique à la base $\mathcal{B}'$

$$P = \begin{pmatrix} 1 & -3 & 5 \\ 1 & 9 & 0 \\ -1 & -12 & 2 \end{pmatrix}$$

répond à la question.

(iii) On a immédiatement, en notant $B = P^{-1}AP$ :

$$B^n = \begin{pmatrix} (-1)^n & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \forall n \geqslant 2.$$

Donc

$$e^{tB} = I + t\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} + \sum_{n=2}^{+\infty} \frac{t^n}{n!}\begin{pmatrix} (-1)^n & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} e^{-t} & 0 & 0 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}.$$

Or $e^{tB} = P^{-1}e^{tA}P$, donc $e^{tA} = Pe^{tB}P^{-1}$. Le calcul de $P^{-1}$ donne

$$P^{-1} = \begin{pmatrix} 2 & -6 & -5 \\ -\frac{2}{9} & \frac{7}{9} & \frac{5}{9} \\ -\frac{1}{3} & \frac{5}{3} & \frac{4}{3} \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 18 & -54 & -45 \\ -2 & 7 & 5 \\ -3 & 15 & 12 \end{pmatrix}.$$

Donc

$$\begin{aligned} e^{tA} &= \begin{pmatrix} 1 & -3 & 5 \\ 1 & 9 & 0 \\ -1 & -12 & 2 \end{pmatrix}\begin{pmatrix} e^{-t} & 0 & 0 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}\frac{1}{9}\begin{pmatrix} 18 & -54 & -45 \\ -2 & 7 & 5 \\ -3 & 15 & 12 \end{pmatrix} \\ &= \frac{1}{9}\begin{pmatrix} 1 & -3 & 5 \\ 1 & 9 & 0 \\ -1 & -12 & 2 \end{pmatrix}\begin{pmatrix} 18e^{-t} & -54e^{-t} & -45e^{-t} \\ -2-3t & 7+15t & 5+12t \\ -3 & 15 & 12 \end{pmatrix} \\ &= \begin{pmatrix} 2e^{-t}-1+t & -6e^{-t}+6-5t & -5e^{-t}+5-4t \\ 2e^{-t}-2-3t & -6e^{-t}+7+15t & -5e^{-t}+5+12t \\ -2e^{-t}+2+4t & 6e^{-t}-6-20t & 5e^{-t}-4-16t \end{pmatrix} \end{aligned}$$

**Exercice 2.9.**

(i) On considère les matrices réelles antisymétriques

$$\mathcal{J}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \mathcal{J}_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Calculer $e^{t\mathcal{J}_1}$ et $e^{t\mathcal{J}_3}$. Quelles transformations géométriques représentent-elles dans la base canonique de $\mathbb{R}^3$?

(ii) Montrer que l'application

$$(\varphi, \theta, \psi) \mapsto e^{\varphi\mathcal{J}_3} e^{\theta\mathcal{J}_1} e^{\psi\mathcal{J}_3} \tag{2.37}$$

de $[0, 2\pi[\times[0, \pi] \times [0, 2\pi[$ dans $SO(3)$ est surjective.

(iii) Expliciter la matrice $e^{\varphi\mathcal{J}_3} e^{\theta\mathcal{J}_1} e^{\psi\mathcal{J}_3}$. En déduire que la restriction de l'application (2.37) à $[0, 2\pi[\times]0, \pi[\times[0, 2\pi[$ est une bijection sur

$$SO(3) \setminus \left(\{e^{t\mathcal{J}_3};\ t \in \mathbb{R}\} \cup \{e^{\pi\mathcal{J}_1} e^{t\mathcal{J}_3};\ t \in \mathbb{R}\}\right).$$

$\varphi, \theta, \psi$ sont appelés les *angles d'Euler* de la matrice $e^{\varphi\mathcal{J}_3} e^{\theta\mathcal{J}_1} e^{\psi\mathcal{J}_3}$.

**Indication.**

(i) On a facilement

$$e^{t\mathcal{J}_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t & -\sin t \\ 0 & \sin t & \cos t \end{pmatrix}, e^{t\mathcal{J}_3} = \begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Les transformations représentées sont les rotations d'angle $t$ autour de l'axe orienté $Ox$ et $Oz$ respectivement.

(ii) Soit $A \in SO(3)$ et $\mathcal{B} = (\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3})$ la base orthonormée canonique de $\mathbb{R}^3$. Posons $\mathbf{f_j} = A\mathbf{e_j}$ pour $j = 1, 2, 3$. $\mathcal{B}' = (\mathbf{f_1}, \mathbf{f_2}, \mathbf{f_3})$ est une base orthonormée directe de $\mathbb{R}^3$.

Si $\mathbf{f_3} = \mathbf{e_3}$, il existe $\psi \in [0, 2\pi[$ tel que $A = e^{\psi\mathcal{J}_3}$, donc $A = e^{0\mathcal{J}_3} e^{0\mathcal{J}_1} e^{\psi\mathcal{J}_3}$. Si $\mathbf{f_3} = -\mathbf{e_3}$, la matrice $B = e^{\pi\mathcal{J}_1} A$ vérifie $B\mathbf{e_3} = \mathbf{e_3}$, et on est ramené au premier cas: il existe $\psi \in [0, 2\pi[$ tel que $B = e^{\psi\mathcal{J}_3}$, donc $A = e^{\pi\mathcal{J}_1} e^{\psi\mathcal{J}_3} = e^{0\mathcal{J}_3} e^{\pi\mathcal{J}_1} e^{\psi\mathcal{J}_3}$.

Supposons donc $\mathbf{f_3} \neq \pm\mathbf{e_3}$. Considérons $A$ comme la matrice de passage $P_{\mathcal{B},\mathcal{B}'}$ de la base canonique $\mathcal{B}$ à la base orthonormée directe $\mathcal{B}' = (\mathbf{f_1}, \mathbf{f_2}, \mathbf{f_3})$. Si $\mathbf{u}$ désigne un vecteur unitaire directeur de la droite intersection des plans $(\mathbf{e_1}, \mathbf{e_2})$ et $(\mathbf{f_1}, \mathbf{f_2})$ orienté convenablement, on a $\mathbf{f_3} = R_{\mathbf{u}}(\theta)\mathbf{e_3}$ pour un $\theta \in ]0, \pi[$. Soit $\mathbf{v}$ tel que la base $\mathcal{C} = (\mathbf{u}, \mathbf{v}, \mathbf{e_3})$ soit orthonormée directe. La matrice de passage de la base canonique à la base $\mathcal{C}$ est $P_{\mathcal{B},\mathcal{C}} = e^{\varphi\mathcal{J}_3}$ pour un $\varphi \in [0, 2\pi[$ unique. Soit $\mathbf{w}$ tel que la base $\mathcal{D} = (\mathbf{u}, \mathbf{w}, \mathbf{f_3})$ soit orthonormée directe. La matrice de passage de la base $\mathcal{C}$ à la base $\mathcal{D}$ est $P_{\mathcal{C},\mathcal{D}} = e^{\theta\mathcal{J}_1}$. Enfin la matrice de passage de la base $\mathcal{D}$ à la base $\mathcal{B}'$ est $P_{\mathcal{D},\mathcal{B}'} = e^{\psi\mathcal{J}_3}$ pour un $\psi \in [0, 2\pi[$ unique. On en déduit que

$$A = P_{\mathcal{B},\mathcal{B}'} = P_{\mathcal{B},\mathcal{C}} P_{\mathcal{C},\mathcal{D}} P_{\mathcal{D},\mathcal{B}'} = e^{\varphi\mathcal{J}_3} e^{\theta\mathcal{J}_1} e^{\psi\mathcal{J}_3}.$$

(iii)

$$e^{\varphi\mathcal{J}_3} e^{\theta\mathcal{J}_1} e^{\psi\mathcal{J}_3} = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \cos\varphi\cos\psi - \sin\varphi\sin\psi\cos\theta & -\cos\varphi\sin\psi - \sin\varphi\cos\psi\cos\theta & \sin\varphi\sin\theta \\ \sin\varphi\cos\psi + \cos\varphi\sin\psi\cos\theta & -\sin\varphi\sin\psi + \cos\varphi\cos\psi\cos\theta & -\cos\varphi\sin\theta \\ \sin\psi\sin\theta & \cos\psi\sin\theta & \cos\theta \end{pmatrix} \tag{2.38}$$

Soit maintenant $A \in SO(3)$ quelconque. On sait d'après (ii) que $A$ s'écrit

$$A = e^{\varphi \mathcal{J}_3} e^{\theta \mathcal{J}_1} e^{\psi \mathcal{J}_3}$$

pour au moins un triplet $(\varphi, \theta, \psi) \in [0, 2\pi[ \times [0, \pi] \times [0, 2\pi[$. D'après la formule (2.38), la troisième colonne de $A$ est $\begin{smallmatrix} \sin\varphi \sin\theta \\ -\cos\varphi \sin\theta \\ \cos\theta \end{smallmatrix}$ et la troisième ligne est $\sin\psi \sin\theta, \cos\psi \sin\theta, \cos\theta$ donc $\cos\theta$ est déterminé de façon unique par $A$, ainsi que $\varphi$ et $\psi$ si $\sin\theta \neq 0$, *i.e.* $\theta \neq 0, \pi$. Cela prouve que l'application $(\varphi, \theta, \psi) \mapsto e^{\varphi \mathcal{J}_3} e^{\theta \mathcal{J}_1} e^{\psi \mathcal{J}_3}$ est injective sur $[0, 2\pi[ \times ]0, \pi[ \times [0, 2\pi[$, et que son image $\Omega$ est l'ensemble des matrices $A \in SO(3)$ pour lesquelles $\sin\theta \neq 0$, *i.e.* $\theta \neq 0, \pi$.

Si $\theta = 0$, on a $A = e^{\varphi \mathcal{J}_3} e^{\psi \mathcal{J}_3} = e^{(\varphi + \psi)\mathcal{J}_3}$ (seule la somme $\varphi + \psi$ est déterminée de façon unique par $A$). Si $\theta = \pi$, on a $A = e^{\varphi \mathcal{J}_3} e^{\pi \mathcal{J}_1} e^{\psi \mathcal{J}_3}$. Or la formule (2.21) donne $e^{\varphi \mathcal{J}_3} e^{\pi \mathcal{J}_1} = e^{\pi \mathcal{J}_1} e^{-\varphi \mathcal{J}_3}$. Donc $A = e^{\pi \mathcal{J}_1} e^{(\psi - \varphi)\mathcal{J}_3}$ (dans ce cas seule la différence $\psi - \varphi$ est déterminée de façon unique par $A$). On constate donc que le complémentaire de $\Omega$ dans $SO(3)$ est bien $\{e^{t\mathcal{J}_3};\ t \in \mathbb{R}\} \cup \{e^{\pi \mathcal{J}_1} e^{t \mathcal{J}_3};\ t \in \mathbb{R}\}$.

**Exercice 2.10.**

On rappelle que toute matrice $A \in M_n(\mathbb{C})$ est semblable à une matrice triangulaire, *i.e.* $A = PTP^{-1}$, $P \in GL(n, \mathbb{C})$, $T$ triangulaire.

(i) Soit $A \in M_n(\mathbb{C})$. Montrer que la trace de $A$ est la somme des valeurs propres de $A$, répétées avec leur multiplicité.

(ii) Montrer que $\det e^A = e^{\operatorname{Tr} A}$.

(iii) Que dire de $e^A$ si $A$ est antisymétrique?

(iv) On suppose dans toute la suite que $A \in M_3(\mathbb{R})$. Montrer que si $A$ est antisymétrique, alors $e^{sA} \in SO(3) \quad \forall s \in \mathbb{R}$.

(v) Montrer que si $e^{sA} \in SO(3) \quad \forall s \in \mathbb{R}$, alors $A$ est antisymétrique.

(vi) Soit $\theta \in \mathbb{R}$ et

$$B = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Trouver une matrice $J$ telle que $B = e^{\theta J}$.

(vii) En déduire que pour tout $A \in SO(3)$, il existe une matrice antisymétrique $X$ telle que $A = e^X$. $X$ est-elle unique?

**Indication.**

(i) Il existe une matrice $T$ triangulaire supérieure et $P \in GL(n, \mathbb{C})$ telles que $A = PTP^{-1}$. Le polynôme caractéristique de $A$ est égal au polynôme caractéristique de $T$. Or si $\lambda_1, \cdots, \lambda_n$ désignent les éléments diagonaux de $T$, le polynôme caractéristique de $T$ est $\prod_{i=1}^n (\lambda_i - X)$. Les valeurs propres de $A$ (répétées avec leur multiplicité) sont donc $\lambda_1, \cdots, \lambda_n$. Or on a $\operatorname{Tr} A = \operatorname{Tr} PTP^{-1} = \operatorname{Tr} T = \sum_{i=1}^n \lambda_i$. D'où le résultat.

(ii) Avec les notations de (i), on a $e^A = e^{PTP^{-1}} = Pe^T P^{-1}$ (Prop. 2.12). Cela implique $\det e^A = \det e^T$. Or comme $T$ est triangulaire supérieure, avec pour éléments diagonaux $\lambda_1, \ldots, \lambda_n$, pour tout $k \in \mathbb{N}$ la matrice $T^k$ est triangulaire supérieure, avec pour éléments diagonaux $\lambda_1^k, \ldots, \lambda_n^k$ donc $e^T$ est triangulaire supérieure, avec pour éléments diagonaux $e^{\lambda_1}, \ldots, e^{\lambda_n}$. Alors $\det e^T = \prod_{i=1}^n e^{\lambda_i} = e^{\sum_{i=1}^n \lambda_i} = e^{\operatorname{Tr} A}$.

(iii) L'application $X \mapsto {}^t A$ de $M_n(\mathbb{C})$ dans lui-même étant continue, on a

$${}^t(e^A) = {}^t\left( \sum_{k=0}^{+\infty} \frac{1}{k!} A^k \right) = \sum_{k=0}^{+\infty} \frac{1}{k!} {}^t(A^k) = \sum_{k=0}^{+\infty} \frac{1}{k!} ({}^t A)^k = e^{{}^t A}.$$

Donc si $A$ est antisymétrique, ${}^t(e^A) = e^{-A} = (e^A)^{-1}$. De plus $\det e^A = e^{\operatorname{Tr} A} = 1$, car les termes diagonaux de $A$ sont nuls puisque $A$ est antisymétrique.
(iv) Si $A$ est réelle, il en est de même de $e^A$. Si $A$ est de plus antisymétrique, il en est de même de $sA$. D'après (iii), on a donc ${}^t(e^{sA}) = (e^{sA})^{-1}$, et $\det e^{sA} = e^{s\operatorname{Tr} A} = 1$. Donc $e^{sA} \in SO(3) \quad \forall s \in \mathbb{R}$.
(v) On a par hypothèse ${}^t(e^{sA}) = (e^{sA})^{-1}$, *i.e.* $e^{s\,{}^tA} = e^{-sA} \quad \forall s \in \mathbb{R}$. Par dérivation, il vient

$$\left[\frac{d}{ds} e^{s\,{}^tA}\right]_{s=0} = \left[\frac{d}{ds} e^{-sA}\right]_{s=0},$$

qui s'écrit

$${}^tA = -A.$$

Donc $A$ est antisymétrique.
(vi) $B = e^{\theta J}$ en posant

$$J = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

(vii) On sait qu'il existe $P \in SO(3)$ et $\theta \in \mathbb{R}$ tels que $A = PBP^{-1}$ (Th.2.6). Or $B = e^{\theta J}$ Donc $A = Pe^{\theta J}P^{-1} = e^{\theta PJP^{-1}} = e^X$ en posant $X = \theta PJP^{-1}$. La matrice $X$ est antisymétrique puisque

$${}^tX = {}^t(\theta PJP^{-1}) = \theta\, {}^t(P^{-1})\, {}^tJ\, {}^tP = -\theta PJP^{-1} = -X$$

car ${}^tP = P^{-1}$ et $J$ est antisymétrique. La matrice $X$ répond à la question. Comme elle dépend de $\theta$ et que $\theta$ n'est défini qu'à $2\pi$ près, il existe une infinité de telles matrices.

**Exercice 2.11.**

Pour $A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in M_2(\mathbb{C})$, on note $A^*$ la matrice conjuguée transposée: $A^* = {}^t\bar{A} = \begin{pmatrix} \bar\alpha & \bar\gamma \\ \bar\beta & \bar\delta \end{pmatrix}$. La matrice $A$ est dite *unitaire* si $AA^* = A^*A = I$; elle est dite *antihermitienne* si $A^* = -A$.

(i) (a) Soit $G \subset M_2(\mathbb{C})$ l'ensemble des matrices unitaires de déterminant 1. Montrer que $G$ est un sous-groupe de $GL(2, \mathbb{C})$.
(b) Déterminer explicitement les éléments de $G$.

(ii) Soit $\mathfrak{g} \subset M_2(\mathbb{C})$ l'ensemble des matrices antihermitiennes de trace nulle.
(a) Montrer que $\mathfrak{g}$ est un sous-espace vectoriel du $\mathbb{R}$-espace vectoriel $M_2(\mathbb{C})$.
(b) Montrer que l'application $\Psi$ définie par

$$\mathbf{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto X_{\mathbf{v}} = \begin{pmatrix} iz & -y+ix \\ y+ix & -iz \end{pmatrix}$$

est un isomorphisme de l'espace vectoriel $\mathbb{R}^3$ sur $\mathfrak{g}$.

(iii) Calculer le déterminant de $X_{\mathbf{v}}$.

(iv) Soit $A \in G$. Montrer que l'application $X \mapsto AXA^{-1}$ est un endomorphisme de $\mathfrak{g}$, que l'on notera $\rho_A$.

(v) Soit $R(A)$ l'endomorphisme $\Psi^{-1} \circ \rho_A \circ \Psi$ de $\mathbb{R}^3$. On identifie $R(A)$ à sa matrice dans la base orthonormée canonique $(\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3})$ de $\mathbb{R}^3$.
(a) Vérifier que $AX_{\mathbf{v}}A^{-1} = X_{R(A)\mathbf{v}}$ $\forall \mathbf{v} \in \mathbb{R}^3$, $\forall A \in G$. En déduire que $R(A) \in O(3)$.

(b) Calculer $R(A)\mathbf{e_i}$ $(1 \leqslant i \leqslant 3)$. En déduire une expression de $R(A)$ pour $A \in G$.
(c) Montrer que $R(A) \in SO(3)$ $\forall A \in G$.
(d) Montrer que l'application $R$ définie par $A \mapsto R(A)$ est un homomorphisme du groupe $G$ dans le groupe $SO(3)$.
(e) Calculer $\operatorname{Ker} R$. (On cherchera les $A \in G$ telles que $AX_{\mathbf{v}} = X_{\mathbf{v}}A$ $\forall \mathbf{v} \in \mathbb{R}^3$).

(vi) Calculer pour $t \in \mathbb{R}$ les matrices

$$R\left(\begin{pmatrix} e^{i\frac{t}{2}} & 0 \\ 0 & e^{-i\frac{t}{2}} \end{pmatrix}\right) \quad , \quad R\left(\begin{pmatrix} \cos\frac{t}{2} & i\sin\frac{t}{2} \\ i\sin\frac{t}{2} & \cos\frac{t}{2} \end{pmatrix}\right).$$

(vii) En déduire que l'image de $R$ est $SO(3)$. (On rappelle que tout élément de $SO(3)$ admet une décomposition avec les angles d'Euler).

(viii) Montrer que $SO(3)$ est isomorphe au groupe quotient $G/\{\pm I\}$.

(ix) Comment peut-on paramétrer $G$ par les angles d'Euler?

**Indication.**

(i)(a) • $G \subset GL(2,\mathbb{C})$ puisqu'une matrice $A$ est unitaire si et seulement si elle est inversible et que son inverse est $A^*$.
• $G$ est stable par la multiplication dans $GL(2,\mathbb{C})$. En effet, si $A, B \in G$, on a $(AB)^* = B^*A^* = B^{-1}A^{-1} = (AB)^{-1}$ donc $AB$ est unitaire. De plus $\det AB = \det A \det B = 1$. Donc $AB \in G$.
• $G$ est stable par passage à l'inverse dans $GL(2,\mathbb{C})$. En effet, si $A \in G$, on a $A^{-1} = A^*$ donc $(A^{-1})^* = A^{**}$. Or $A^{**} = A$. Donc $(A^{-1})^* = A = (A^{-1})^{-1}$ et $A^{-1}$ est unitaire. De plus $\det(A^{-1}) = (\det A)^{-1} = 1$. Donc $A^{-1} \in G$.
(b) Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{C})$. On a $A \in G$ si et seulement si $A^* = A^{-1}$ et $\det A = 1$. Compte tenu de $\det A = 1$, l'équation $A^* = A^{-1}$ s'écrit :

$$\begin{pmatrix} \bar\alpha & \bar\gamma \\ \bar\beta & \bar\delta \end{pmatrix} = \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix}.$$

Cette dernière équation équivaut à $\delta = \bar\alpha$ et $\gamma = -\bar\beta$ donc

$$A = \begin{pmatrix} \alpha & \beta \\ -\bar\beta & \bar\alpha \end{pmatrix}$$

avec $|\alpha|^2 + |\beta|^2 = 1$. Réciproquement, toute matrice de ce type est unitaire de déterminant 1. On a donc :

$$G = \left\{ \begin{pmatrix} \alpha & \beta \\ -\bar\beta & \bar\alpha \end{pmatrix} ;\ \alpha, \beta \in \mathbb{C},\ |\alpha|^2 + |\beta|^2 = 1 \right\}.$$

$G$ n'est autre que l'ensemble de quaternions de norme 1 (voir Ex. 1.20).
(ii)(a) Soient $A, B \in \mathfrak{g}$ et $\lambda \in \mathbb{R}$. On a $(A+B)^* = A^* + B^* = -A - B = -(A+B)$, $(\lambda A)^* = \lambda A^* = -\lambda A$ et $\operatorname{Tr}(A+B) = \operatorname{Tr} A + \operatorname{Tr} B = 0$, $\operatorname{Tr} \lambda A = \lambda \operatorname{Tr} A = 0$. Donc $A + B \in \mathfrak{g}$ et $\lambda A \in \mathfrak{g}$.
(b) Il est immédiat que $\Psi$ est bien une application de $\mathbb{R}^3$ dans $\mathfrak{g}$, et qu'elle est linéaire et injective. Montrons qu'elle est surjective. Soit $X = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathfrak{g}$. On a par définition $X^* = -X$ et $\operatorname{Tr} X = 0$ donc $a$ est imaginaire pur, $d = -a$, et $c = -\bar b$. Il existe donc $x, y, z \in \mathbb{R}$ tels que $a = iz$, $b = -y + ix$ et alors

$$X = \begin{pmatrix} iz & -y+ix \\ y+ix & -iz \end{pmatrix} = X_{\mathbf{v}} = \Psi(\mathbf{v}) \quad \text{avec} \quad \mathbf{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

(iii) $\det X_{\mathbf{v}} = x^2 + y^2 + z^2 = \|\mathbf{v}\|^2$.
(iv) Pour $A \in G$, on a $A^{-1} = A^*$. Donc

$$(AXA^{-1})^* = (AXA^*)^* = A^{**}X^*A^* = AX^*A^* = -AXA^* = -AXA^{-1} \ \forall X \in \mathfrak{g}.$$

Comme de plus $\operatorname{Tr} AXA^{-1} = \operatorname{Tr} X = 0 \ \forall X \in \mathfrak{g}$, on a $AXA^{-1} \in \mathfrak{g} \ \forall X \in \mathfrak{g}$, *i.e.* l'application $\varrho_A : X \mapsto AXA^{-1}$ est une application de $\mathfrak{g}$ dans lui-même. Par ailleurs, $A(X+Y)A^{-1} = AXA^{-1} + AYA^{-1} \ \forall X, Y \in \mathfrak{g}$, et $A(\lambda X)A^{-1} = \lambda AXA^{-1}$ $\forall X \in \mathfrak{g}$, $\forall \lambda \in \mathbb{R}$. Donc $\varrho_A$ est un endomorphisme de $\mathfrak{g}$.
(v)(a) Pour tout $\mathbf{v} \in \mathbb{R}^3$ on a

$$R(A)\mathbf{v} = \left(\Psi^{-1} \circ \rho_A \circ \Psi\right)(\mathbf{v}) = \Psi^{-1}\left(AX_{\mathbf{v}}A^{-1}\right),$$

donc $\Psi(R(A)\mathbf{v}) = AX_{\mathbf{v}}A^{-1}$, *i.e.* $X_{R(A)\mathbf{v}} = AX_{\mathbf{v}}A^{-1}$.

On en déduit que $\det X_{R(A)\mathbf{v}} = \det AX_{\mathbf{v}}A^{-1} = \det X_{\mathbf{v}}$, ce qui s'écrit d'après la question (iii) $\|R(A)\mathbf{v}\|^2 = \|\mathbf{v}\|^2 \ \forall \mathbf{v} \in \mathbb{R}^3$, donc $R(A) \in O(3)$.
(b) Soit

$$A = \begin{pmatrix} \alpha & \beta \\ -\bar\beta & \bar\alpha \end{pmatrix} \in G\,,\ \alpha, \beta \in \mathbb{C},\ |\alpha|^2 + |\beta|^2 = 1\,.$$

On a :

$$X_{R(A)\mathbf{e_1}} = AX_{\mathbf{e_1}}A^{-1} = \begin{pmatrix} \alpha & \beta \\ -\bar\beta & \bar\alpha \end{pmatrix}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}\begin{pmatrix} \bar\alpha & -\beta \\ \bar\beta & \alpha \end{pmatrix} = \begin{pmatrix} i(\alpha\bar\beta + \bar\alpha\beta) & i(\alpha^2 - \beta^2) \\ i(\bar\alpha^2 - \bar\beta^2) & -i(\alpha\bar\beta + \bar\alpha\beta) \end{pmatrix}.$$

Donc

$$R(A)\mathbf{e_1} = \begin{pmatrix} \mathfrak{Re}\,(\alpha^2 - \beta^2) \\ \mathfrak{Im}\,(\alpha^2 - \beta^2) \\ 2\mathfrak{Re}\,(\alpha\bar\beta) \end{pmatrix}.$$

De même :

$$X_{R(A)\mathbf{e_2}} = AX_{\mathbf{e_2}}A^{-1} = \begin{pmatrix} \alpha & \beta \\ -\bar\beta & \bar\alpha \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \bar\alpha & -\beta \\ \bar\beta & \alpha \end{pmatrix} = \begin{pmatrix} -\alpha\bar\beta + \bar\alpha\beta & -\alpha^2 - \beta^2 \\ \bar\alpha^2 + \bar\beta^2 & \alpha\bar\beta - \bar\alpha\beta \end{pmatrix},$$

donc

$$R(A)\mathbf{e_2} = \begin{pmatrix} -\mathfrak{Im}\,(\alpha^2 + \beta^2) \\ \mathfrak{Re}\,(\alpha^2 + \beta^2) \\ -2\mathfrak{Im}\,(\alpha\bar\beta) \end{pmatrix};$$

$$X_{R(A)\mathbf{e_3}} = AX_{\mathbf{e_3}}A^{-1} = \begin{pmatrix} \alpha & \beta \\ -\bar\beta & \bar\alpha \end{pmatrix}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}\begin{pmatrix} \bar\alpha & -\beta \\ \bar\beta & \alpha \end{pmatrix}$$
$$= \begin{pmatrix} i(|\alpha|^2 - |\beta|^2) & -2i\alpha\beta \\ -2i\bar\alpha\bar\beta & -i(|\alpha|^2 - |\beta|^2) \end{pmatrix},$$

donc

$$R(A)\mathbf{e_3} = \begin{pmatrix} -2\mathfrak{Re}\,(\alpha\beta) \\ -2\mathfrak{Im}\,(\alpha\beta) \\ |\alpha|^2 - |\beta|^2 \end{pmatrix}.$$

Cela donne

$$R(A) = \begin{pmatrix} \mathfrak{Re}\,(\alpha^2 - \beta^2) & -\mathfrak{Im}\,(\alpha^2 + \beta^2) & -2\mathfrak{Re}\,(\alpha\beta) \\ \mathfrak{Im}\,(\alpha^2 - \beta^2) & \mathfrak{Re}\,(\alpha^2 + \beta^2) & -2\mathfrak{Im}\,(\alpha\beta) \\ 2\mathfrak{Re}\,(\alpha\bar{\beta}) & -2\mathfrak{Im}\,(\alpha\bar{\beta}) & |\alpha|^2 - |\beta|^2 \end{pmatrix}. \tag{2.39}$$

$R(A)$ s'écrit aussi en notant $\alpha = u + iv$ , $\beta = w + it$ avec $u, v, w, t \in \mathbb{R}$, $u^2 + v^2 + w^2 + t^2 = 1$ :

$$R(A) = \begin{pmatrix} u^2 - v^2 - w^2 + t^2 & -2(uv + wt) & -2(uw - vt) \\ 2(uv - wt) & u^2 - v^2 + w^2 - t^2 & -2(ut + vw) \\ 2(uw + vt) & -2(-ut + vw) & u^2 + v^2 - w^2 - t^2 \end{pmatrix}. \tag{2.40}$$

(c) Il s'agit de montrer que $\det R(A) = 1\ \forall A \in G$. Soit donc $A = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \in G$. Comme $|\alpha|^2 + |\beta|^2 = 1$, il existe $\lambda \in \mathbb{R}$ tel que $|\alpha| = \cos\lambda$ , $|\beta| = \sin\lambda$. On a alors $\alpha = e^{i\mu}\cos\lambda$ , $\beta = e^{i\nu}\sin\lambda$ avec $\mu, \nu \in \mathbb{R}$. Soit maintenant pour $s \in [0,1]$

$$A_s = \begin{pmatrix} e^{is\mu}\cos(s\lambda) & e^{is\nu}\sin(s\lambda) \\ -e^{-is\nu}\sin(s\lambda) & e^{-is\mu}\cos(s\lambda) \end{pmatrix}.$$

On a $A_s \in G$ pour tout $s$, et $A_0 = I$, $A_1 = A$. D'après la formule (2.39) appliquée à $A_s$, la fonction $f(s) = \det R(A_s)$ est une fonction continue. Comme $R(A_s) \in O(3)$, la fonction $f$ est à valeurs dans $\{-1, 1\}$. Comme elle est continue, on a d'après le Théorème des valeurs intermédiaires $f(s) = -1\,\forall s \in [0,1]$ ou $f(s) = 1\,\forall s \in [0,1]$. Mais $f(0) = \det R(A_0) = \det R(I) = \det I = 1$. Donc $f(s) = 1\,\forall s \in [0,1]$ et ainsi $\det R(A) = \det R(A_1) = f(1) = 1$.
(d) D'après (c), $R$ est bien une application de $G$ dans $SO(3)$. Il reste à montrer que c'est un homomorphisme de groupes. Pour tous $A, B \in G$ et tout $\mathbf{v} \in \mathbb{R}^3$, on a

$$X_{R(AB)\mathbf{v}} = ABX_{\mathbf{v}}(AB)^{-1} = ABX_{\mathbf{v}}B^{-1}A^{-1} = AX_{R(B)\mathbf{v}}A^{-1} = X_{R(A)R(B)\mathbf{v}}$$

donc $R(AB)\mathbf{v} = R(A)R(B)\mathbf{v}$. Comme $\mathbf{v} \in \mathbb{R}^3$ est arbitraire, $R(AB) = R(A)R(B)$. $A, B \in G$ étant arbitraires, $R$ est un homomorphisme de $G$ dans $SO(3)$.
(e) Soit

$$A = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \in G\,,\ \alpha, \beta \in \mathbb{C},\ |\alpha|^2 + |\beta|^2 = 1\,.$$

On a $A \in \operatorname{Ker} R \Leftrightarrow R(A) = I$. Supposons $R(A) = I$. La formule (2.39) donne immédiatement $\alpha^2 - \beta^2 = 1$ et $\alpha^2 + \beta^2 = 1$. Donc $\beta = 0$ et $\alpha^2 = 1$, *i.e.* $\alpha = \pm 1$. Ainsi $A = \pm I$. Réciproquement, si $A = \pm I$, on a bien $A \in G$ et la formule (2.39) donne $R(A) = I$. Donc $\operatorname{Ker} R = \{\pm I\}$.

(vi) Pour $A = \begin{pmatrix} e^{i\frac{t}{2}} & 0 \\ 0 & e^{-i\frac{t}{2}} \end{pmatrix}$, la formule (2.39) donne

$$R(A) = \begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix} = e^{t\mathcal{J}_3} \quad \text{avec} \quad \mathcal{J}_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Pour $B = \begin{pmatrix} \cos\frac{t}{2} & i\sin\frac{t}{2} \\ i\sin\frac{t}{2} & \cos\frac{t}{2} \end{pmatrix}$, la formule donne :

$$R(B) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t & -\sin t \\ 0 & \sin t & \cos t \end{pmatrix} = e^{t\mathcal{J}_1} \quad \text{avec} \quad \mathcal{J}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}.$$

A noter qu'on a de même

$$R\left(\begin{pmatrix} \cos\frac{t}{2} & -\sin\frac{t}{2} \\ \sin\frac{t}{2} & \cos\frac{t}{2} \end{pmatrix}\right) = \begin{pmatrix} \cos t & 0 & -\sin t \\ 0 & 1 & 0 \\ \sin t & 0 & \cos t \end{pmatrix} = e^{t\mathcal{J}_2} \quad \text{avec} \quad \mathcal{J}_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Remarquons que

$$\begin{pmatrix} e^{i\frac{t}{2}} & 0 \\ 0 & e^{-i\frac{t}{2}} \end{pmatrix} = e^{t\mathcal{X}_3}, \quad \begin{pmatrix} \cos\frac{t}{2} & i\sin\frac{t}{2} \\ i\sin\frac{t}{2} & \cos\frac{t}{2} \end{pmatrix} = e^{t\mathcal{X}_1}, \quad \begin{pmatrix} \cos\frac{t}{2} & -\sin\frac{t}{2} \\ \sin\frac{t}{2} & \cos\frac{t}{2} \end{pmatrix} = e^{t\mathcal{X}_2}$$

avec

$$\mathcal{X}_3 = \frac{1}{2}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad \mathcal{X}_1 = \frac{1}{2}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad \mathcal{X}_2 = \frac{1}{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Alors on a $R(e^{t\mathcal{X}_k}) = e^{t\mathcal{J}_k}$ $\forall k$ $1 \leqslant k \leqslant 3$.
(vii) Soit $M \in SO(3)$. On sait (voir Ex. 2.9) qu'il existe $\varphi, \theta, \psi$ ($\varphi, \psi \in [0, 2\pi[$ ; $\theta \in [0, \pi]$) tels que $M = e^{\varphi\mathcal{J}_3}e^{\theta\mathcal{J}_1}e^{\psi\mathcal{J}_3}$. D'après (vi), cela s'écrit :

$$M = R\left(\begin{pmatrix} e^{i\frac{\varphi}{2}} & 0 \\ 0 & e^{-i\frac{\varphi}{2}} \end{pmatrix}\right) R\left(\begin{pmatrix} \cos\frac{\theta}{2} & i\sin\frac{\theta}{2} \\ i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}\right) R\left(\begin{pmatrix} e^{i\frac{\psi}{2}} & 0 \\ 0 & e^{-i\frac{\psi}{2}} \end{pmatrix}\right) = R(A)$$

avec

$$A = \begin{pmatrix} e^{i\frac{\varphi}{2}} & 0 \\ 0 & e^{-i\frac{\varphi}{2}} \end{pmatrix} \begin{pmatrix} \cos\frac{\theta}{2} & i\sin\frac{\theta}{2} \\ i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix} \begin{pmatrix} e^{i\frac{\psi}{2}} & 0 \\ 0 & e^{-i\frac{\psi}{2}} \end{pmatrix} = e^{\varphi\mathcal{X}_3}e^{\theta\mathcal{X}_1}e^{\psi\mathcal{X}_3} \in G.$$

$M \in SO(3)$ étant arbitraire, cela montre que $R$ est surjective. Noter que

$$A = \begin{pmatrix} e^{i\frac{\varphi+\psi}{2}}\cos\frac{\theta}{2} & ie^{i\frac{\varphi-\psi}{2}}\sin\frac{\theta}{2} \\ ie^{-i\frac{\varphi-\psi}{2}}\sin\frac{\theta}{2} & e^{-i\frac{\varphi+\psi}{2}}\cos\frac{\theta}{2} \end{pmatrix}. \tag{2.41}$$

(viii) $R : G \to SO(3)$ est un homorphisme surjectif de noyau $\operatorname{Ker} R = \{\pm I\}$. Sa décomposition canonique donne donc un isomorphisme $\tilde{R} : G/\{\pm I\} \to SO(3)$.

(ix) Soit $A \in G$. Il existe $\theta, \varphi, \psi$ $0 \leqslant \theta \leqslant \pi$ , $0 \leqslant \varphi < 2\pi$ , $0 \leqslant \psi < 2\pi$ tels que $R(A) = e^{\varphi \mathcal{J}_3} e^{\theta \mathcal{J}_1} e^{\psi \mathcal{J}_3}$. On a vu qu'alors $R(e^{\varphi \mathcal{X}_3} e^{\theta \mathcal{X}_1} e^{\psi \mathcal{X}_3}) = R(A)$. Or pour $A, A' \in G$, $R(A) = R(A') \Leftrightarrow A = \pm A'$ puisque $\operatorname{Ker} R = \{\pm I\}$. Donc $A = \pm e^{\varphi \mathcal{X}_3} e^{\theta \mathcal{X}_1} e^{\psi \mathcal{X}_3}$. On ramène le signe $-$ au signe $+$ en changeant $\psi$ en $\psi - 2\pi$. On voit donc qu'il existe $\theta, \varphi, \psi$,

$$0 \leqslant \theta \leqslant \pi \,,\, 0 \leqslant \varphi < 2\pi \,,\, -2\pi \leqslant \psi < 2\pi \tag{2.42}$$

tels que $A = e^{\varphi \mathcal{X}_3} e^{\theta \mathcal{X}_1} e^{\psi \mathcal{X}_3}$, *i.e.* $A$ est donné par la formule (2.41) avec les paramètres vérifiant (2.42). Cela définit un paramétrage de $G$. D'après les propriétés du paramétrage de $SO(3)$ par les angles d'Euler, ce paramétrage est injectif pour $0 < \theta < \pi$.

On peut aussi raisonner de la façon suivante. Soit

$$A = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \in G \,,\ \alpha, \beta \in \mathbb{C},\ |\alpha|^2 + |\beta|^2 = 1 \,.$$

Il existe $\theta$ $0 \leqslant \theta \leqslant \pi$ unique tel que $|\alpha| = \cos \frac{\theta}{2}$, $|\beta| = \sin \frac{\theta}{2}$. Il existe $\mu, \xi$ $0 \leqslant \mu < 2\pi$, $0 \leqslant \xi < 2\pi$, uniques tels que $\alpha = |\alpha| e^{i\mu}$ , $\beta = |\beta|\, i\, e^{i\xi}$. En écrivant $\mu = \frac{\varphi+\psi}{2}$ , $\xi = \frac{\varphi-\psi}{2}$, il existe $\varphi, \psi$ uniques tels que $0 \leqslant \varphi < 4\pi$ , $-2\pi < \psi < 2\pi$ vérifiant

$$\alpha = |\alpha|\, e^{i\frac{\varphi+\psi}{2}} \,,\ \beta = |\beta|\, i e^{i\frac{\varphi-\psi}{2}} \,. \tag{2.43}$$

En changeant, si $2\pi \leqslant \varphi < 4\pi$, $\varphi$ en $\varphi - 2\pi$ et simultanément $\psi$ en $\psi - 2\pi$ (*resp.* $\psi + 2\pi$) si $0 \leqslant \psi < 2\pi$ (*resp.* $-2\pi < \psi < 0$), on voit qu'il existe $\varphi, \psi$ tels que $0 \leqslant \varphi < 2\pi$ , $-2\pi \leqslant \psi < 2\pi$ vérifiant (2.43). Alors on a la formule (2.41) pour $A$ avec les paramètres vérifiant (2.42). Si $0 < \theta < \pi$, *i.e.* $\alpha\beta \neq 0$, $\varphi$ et $\psi$, $0 \leqslant \varphi < 2\pi$ , $-2\pi \leqslant \psi < 2\pi$ sont uniques. En effet, supposons $e^{i\frac{\varphi+\psi}{2}} = e^{i\frac{\varphi'+\psi'}{2}}$ et $e^{i\frac{\varphi-\psi}{2}} = e^{i\frac{\varphi'-\psi'}{2}}$ avec $\varphi, \varphi', \psi, \psi'$ tels que $0 \leqslant \varphi, \varphi' < 2\pi$ , $-2\pi \leqslant \psi, \psi' < 2\pi$. Par multiplication, on aurait $e^{i\varphi} = e^{i\varphi'}$, donc $\varphi = \varphi'$. Alors $e^{i\frac{\psi}{2}} = e^{i\frac{\psi'}{2}}$, donc $\psi = \psi'$.

# Chapitre 3

# Actions de groupes.

## 3.1 Actions de groupe.

### 3.1.1 Définition.

**Définition 3. 1.** *Soit $G$ un groupe et $X$ un ensemble. $G$ est noté multiplicativement et $e$ désigne son élément neutre. On appelle action de $G$ sur $X$ une application $G \times X \to X$ notée $(g, x) \mapsto g \cdot x$ vérifiant :*

*(i)* $e \cdot x = x \quad \forall x \in X$

*(ii)* $g_1 \cdot (g_2 \cdot x) = (g_1 g_2) \cdot x \quad \forall\, g_1, g_2 \in G,\ \forall\, x \in X.$

### 3.1.2 Exemples.

(i) L'action naturelle de $G = GL(n, \mathbb{R})$ sur $X = \mathbb{R}^n$ définie par :

$$(A, x) \mapsto A \cdot x = Ax \quad \forall\, A \in G,\ \forall\, x \in X.$$

(ii) L'action naturelle de $G = O(n)$ sur la sphère unité $X = \mathbb{S}_n = \{\mathbf{x} \in \mathbb{R}^n;\ \|\mathbf{x}\| = 1\}$ de $\mathbb{R}^n$ définie par

$$(A, x) \mapsto A \cdot x = Ax \quad \forall\, A \in G,\ \forall\, x \in X.$$

(iii) L'action naturelle de $G = \mathcal{S}_n$ sur $X = \{1, \ldots, n\}$ définie par :

$$(\sigma, x) \mapsto \sigma \cdot x = \sigma(x) \quad \forall\, \sigma \in G,\ \forall\, x \in X.$$

(iv) L'action d'un groupe quelconque $G$ sur lui-même par *translation à gauche* définie par :

$$(g, x) \mapsto g \cdot x = L_g(x) = gx \quad \forall\, g \in G,\ \forall\, x \in G.$$

(v) L'action d'un groupe quelconque $G$ sur lui-même par *"translation à droite"* définie par :

$$(g, x) \mapsto g \cdot x = R_{g^{-1}}(x) = xg^{-1} \quad \forall\, g \in G,\ \forall\, x \in G.$$

(vi) L'action d'un groupe quelconque $G$ sur lui-même par *automorphismes intérieurs* définie par :

$$(g, x) \mapsto g \cdot x = \mathrm{Int}\,(g)(x) = gxg^{-1} \quad \forall\, g \in G,\ \forall\, x \in G.$$

(vii) L'action d'un groupe quelconque $G$ sur un sous-groupe distingué $H$ de $G$ par *automorphismes intérieurs* définie par :

$$(g, x) \mapsto g \cdot x = gxg^{-1} \quad \forall\, g \in G,\ \forall\, x \in H.$$

(viii) L'action d'un groupe quelconque $G$ sur l'espace homogène $G/H$, où $H$ est un sous-groupe de $G$, par *translation à gauche* définie par :

$$(g, [x]) \mapsto g \cdot [x] = [gx] \quad \forall\, g \in G,\ \forall\, [x] \in G/H.$$

Cette action est bien définie puisque l'on a pour tout $g \in G$

$$x \equiv x' \mod H \Rightarrow gx \equiv gx' \mod H.$$

### 3.1.3 Stabilisateur et orbite.

**Proposition 3. 1.** *Soit $X$ un ensemble sur lequel agit un groupe $G$. Pour tout point $x \in X$, $S_x = \{g \in G;\ g \cdot x = x\}$ est un sous-groupe de $G$.*

**Démonstration.**
On a $S_x \subset G$ et $e \in S_x$. Si $g, g' \in S_x$,

$$(gg') \cdot x = g \cdot (g' \cdot x) = g \cdot x = x$$

donc $gg' \in S_x$. Enfin, pour $g \in S_x$, l'équation $x = g \cdot x$ donne par action de $g^{-1}$ :

$$g^{-1} \cdot x = g^{-1} \cdot (g \cdot x) = (g^{-1}g) \cdot x = e \cdot x = x$$

donc $g^{-1} \in S_x$. □

**Définition 3. 2.** *Soit $X$ un ensemble sur lequel agit un groupe $G$. Pour tout point $x \in X$, le sous-groupe $S_x = \{g \in G;\ g \cdot x = x\}$ de $G$ est appelé le stabilisateur du point $x \in X$.*

**Définition 3. 3.** *Soit $X$ un ensemble sur lequel agit un groupe $G$. Pour tout point $x \in X$, le sous-ensemble $\mathcal{O}(x) = \{y \in X;\ \exists\, g \in G;\ y = g \cdot x\}$ de $X$ est appelé l'orbite du point $x \in X$. On note aussi $\mathcal{O}(x) = G \cdot x$.*

**Remarque.** La relation $\mathcal{R}$ définie sur $X$ par

$$x\,\mathcal{R}\,y \;\Leftrightarrow\; y \in \mathcal{O}(x)$$

est une relation d'équivalence sur $X$ :
• elle est réflexive puisque pour tout $x \in X$ on a $x = e \cdot x \in \mathcal{O}(x)$;
• elle est symétrique puisque pour tous $x, y \in X$ tels que $y \in \mathcal{O}(x)$, il existe $a \in G$ tel que $y = a \cdot x$ et alors $x = a^{-1} \cdot y$ donc $x \in \mathcal{O}(y)$;
• elle est transitive puisque pour tous $x, y, z \in X$ tels que $y \in \mathcal{O}(x)$ et $z \in \mathcal{O}(y)$, il existe $a, b \in G$ tel que $y = a \cdot x$ et $z = b \cdot y$ et alors $z = b \cdot (a \cdot x) = (ba) \cdot x \in \mathcal{O}(x)$.

La classe d'équivalence d'un élément $x \in X$ est l'orbite $\mathcal{O}(x)$. En particulier, les orbites forment une partition de $X$, et deux orbites sont soit disjointes soit égales : si $y \in \mathcal{O}(x)$, alors $\mathcal{O}(y) = \mathcal{O}(x)$.

**Théorème 3. 1.** *Soit $X$ un ensemble sur lequel agit un groupe $G$, et $x$ un point de $X$. Il existe une bijection canonique de l'espace homogène $G/S_x$ sur l'orbite $\mathcal{O}(x)$.*

**Démonstration.**

Soit $\varphi_x : G \to \mathcal{O}(x) = G \cdot x$ l'application définie par $\varphi_x(g) = g \cdot x \quad \forall\, g \in G$. Par définition de l'orbite $\mathcal{O}(x) = G \cdot x$, l'application $\varphi_x$ est surjective. On a de plus pour tous $g_1, g_2 \in G$ :

$$\begin{aligned} \varphi_x(g_1) = \varphi_x(g_2) &\Leftrightarrow g_1 \cdot x = g_2 \cdot x \\ &\Leftrightarrow {g_1}^{-1} g_2 \cdot x = x \\ &\Leftrightarrow {g_1}^{-1} g_2 \in S_x\,. \end{aligned}$$

$\varphi_x$ définit donc par passage au quotient une application $\tilde{\varphi}_x : G/S_x \to \mathcal{O}(x)$ par la formule

$$\tilde{\varphi}_x([g]) = \varphi_x(g) \quad \forall\, g \in G,$$

puisque si $g_1, g_2 \in G$ sont deux représentants d'une même classe $[g_1] = [g_2] \in G/S_x$, on a $\varphi_x(g_1) = \varphi_x(g_2)$. L'application $\tilde{\varphi}_x$ est surjective puisque $\varphi_x$ est surjective : pour tout $y \in \mathcal{O}(x)$, il existe $g \in G$ tel que $y = g \cdot x = \varphi_x(g) = \tilde{\varphi}_x([g])$. L'application $\tilde{\varphi}_x$ est injective par définition, puisque pour tous $g_1, g_2 \in G$,

$$\begin{aligned} \tilde{\varphi}_x([g_1]) = \tilde{\varphi}_x([g_2]) &\Leftrightarrow \varphi_x(g_1) = \varphi_x(g_2) \\ &\Leftrightarrow {g_1}^{-1} g_2 \in S_x \\ &\Leftrightarrow g_1 \equiv g_2 \pmod{S_x} \\ &\Leftrightarrow [g_1] = [g_2] \end{aligned}$$

L'application $\tilde{\varphi}_x$ est donc une bijection canonique de l'espace homogène $G/S_x$ sur l'orbite $\mathcal{O}(x)$. □

### 3.1.4 Équation des classes.

**Théorème 3. 2.** *Soit $G$ un groupe fini, $Z(G)$ son centre et $\Omega_i$ $(1 \leqslant i \leqslant k)$ les classes de conjugaison des éléments de $G \setminus Z(G)$. Alors*

$$|G| = |Z(G)| + \sum_{i=1}^{k} |\Omega_i|. \tag{3.1}$$

**Démonstration.**

Considérons l'action de $G$ sur lui même par automorphismes intérieurs :

$$(g, x) \mapsto g \cdot x = \operatorname{Int}(g)(x) = gxg^{-1} \quad \forall\, g \in G,\ \forall\, x \in G.$$

L'orbite $\mathcal{O}(x)$ d'un élément $x \in G$ est la classe de conjugaison de $x$ (Déf. 1.9). On a $\mathcal{O}(x) = \{x\}$ si et seulement si $x \in Z(G)$. Le nombre d'orbites réduites à un point est donc le cardinal de $Z(G)$. Les orbites qui ne sont pas réduites à un point sont celles des éléments de $G \setminus Z(G)$, *i.e.* les $\Omega_i$ $(1 \leqslant i \leqslant k)$. Comme les orbites forment une partition de $G$, l'équation (3.1) en résulte. □

### 3.1.5 Nombre d'orbites.

**Théorème 3. 3 (Formule de Burnside).** *Soit $X$ un ensemble fini sur lequel agit un groupe fini $G$. Le nombre $N$ d'orbites de $X$ sous l'action de $G$ est donné par la formule*

$$N = \frac{1}{|G|} \sum_{g \in G} |X^g| \tag{3.2}$$

*où pour tout $g \in G$, on pose $X^g = \{x \in X;\ g \cdot x = x\}$.*

**Démonstration.**
Soit

$$\mathfrak{X} = \{(g,x) \in G \times X;\ g \cdot x = x\}.$$

On a

$$\mathfrak{X} = \bigcup_{g \in G} (\{g\} \times \{x \in X;\ g \cdot x = x\}) = \bigcup_{g \in G} (\{g\} \times X^g), \tag{3.3}$$

$$\mathfrak{X} = \bigcup_{x \in X} (\{g \in G;\ g \cdot x = x\} \times \{x\}) = \bigcup_{x \in X} (S_x \times \{x\}). \tag{3.4}$$

On calcule le cardinal de $\mathfrak{X}$ de deux façons différentes en utilisant les équations (3.3) et (3.4). D'après (3.3), on a :

$$|\mathfrak{X}| = \sum_{g \in G} |X^g|. \tag{3.5}$$

D'après (3.4), on a, en désignant par $N$ le nombre d'orbites distinctes et par $\mathcal{O}_1, \cdots, \mathcal{O}_N$ les orbites distinctes,

$$|\mathfrak{X}| = \sum_{x \in X} |S_x| = \sum_{i=1}^{N} \left( \sum_{x \in \mathcal{O}_i} |S_x| \right). \tag{3.6}$$

Mais pour tout $i = 1, \cdots, N$ et tout $x \in \mathcal{O}_i$ on a

$$\mathcal{O}(x) = \mathcal{O}_i$$

donc d'après le Théorème 3.1,

$$|S_x| = \frac{|G|}{|\mathcal{O}(x)|} = \frac{|G|}{|\mathcal{O}_i|}.$$

L'équation (3.6) s'écrit alors :

$$|\mathfrak{X}| = \sum_{i=1}^{N} \sum_{x \in \mathcal{O}_i} \frac{|G|}{|\mathcal{O}_i|} = \sum_{i=1}^{N} |G| = N|G|. \tag{3.7}$$

Les équations (3.5) et (3.7) donnent alors :

$$N = \frac{1}{|G|} \sum_{g \in G} |X^g|.$$

□

## 3.2 Théorème de Cauchy.

**Théorème 3. 4.** *Soit $G$ un groupe fini d'ordre $n$ et $p$ un diviseur premier de $n$. Il existe un sous-groupe $H$ de $G$ d'ordre $p$.*

**Démonstration.**

On va montrer qu'il existe un élément $x \in G$ d'ordre $p$. Alors $H = \langle x \rangle$ est un sous-groupe d'ordre $p$.

Soit
$$X = \{(x_1, x_2, \dots, x_p) \in G^p;\ x_1 x_2 \cdots x_p = e\}$$
où $e$ est l'élément neutre de $G$ et $G^p$ désigne l'ensemble produit cartésien $\underbrace{G \times \cdots \times G}_{p \text{ fois}}$.

Pour $x \in G$, on a $x^p = e$ si et seulement si $(x, x, \dots, x) \in X$. Comme $p$ est premier, $x$ est donc d'ordre $p$ si et seulement si $x \neq e$ et $(x, x, \dots, x) \in X$. Nous allons montrer qu'il existe un tel $x \neq e$.

On a $|X| = n^{p-1}$ puisque $x_1, x_2, \dots, x_{p-1}$ peuvent être choisis arbitrairement, $x_p$ étant déterminé par $x_p = (x_1 x_2 \cdots x_{p-1})^{-1}$.

Posons pour $(x_1, x_2, \dots, x_p) \in G^p$ et $s \in \mathcal{S}_p$ :
$$s \cdot (x_1, \dots, x_p) = (x_{s^{-1}(1)}, \dots, x_{s^{-1}(p)}). \tag{3.8}$$

La formule (3.8) définit une application
$$\mathcal{S}_p \times G^p \longrightarrow G^p. \tag{3.9}$$

L'application (3.9) est une action de $\mathcal{S}_p$ sur $G^p$. En effet, pour $s = Id$, on a $Id \cdot (x_1, \dots, x_p) = (x_1, \dots, x_p)$, et pour $s, t \in \mathcal{S}_p$ on a
$$\begin{aligned}
s \cdot (t \cdot (x_1, \dots, x_p)) &= s \cdot (x_{t^{-1}(1)}, \dots, x_{t^{-1}(p)}) \\
&= s \cdot (y_1, \dots, y_p) \text{ (en posant } y_j = x_{t^{-1}(j)}\ \forall\, j \text{ )} \\
&= (y_{s^{-1}(1)}, \dots, y_{s^{-1}(p)}) \\
&= (x_{t^{-1}(s^{-1}(1))}, \dots, x_{t^{-1}(s^{-1}(p))}) \\
&= (x_{(st)^{-1}(1)}, \dots, x_{(st)^{-1}(p)}) \\
&= (st) \cdot (x_1, \dots, x_p).
\end{aligned}$$

Soit maintenant $\sigma$ la permutation circulaire $(1, 2, \dots, p) \in \mathcal{S}_p$. Le sous-groupe $\langle \sigma \rangle$ engendré par $\sigma$ dans $\mathcal{S}_p$ est cyclique d'ordre $p$. Or on a
$$\sigma \cdot (x_1, \dots, x_p) \in X \quad \forall\, (x_1, \dots, x_p) \in X. \tag{3.10}$$

En effet, si $(x_1, x_2, \dots, x_p) \in X$, $x_1 x_2 \dots x_{p-1} x_p = e$ donc $x_p = (x_1 x_2 \dots x_{p-1})^{-1}$ et alors $x_p x_1 x_2 \dots x_{p-1} = e$, *i.e.* $\sigma \cdot (x_1, x_2, \dots, x_p) = (x_p, x_1, \dots, x_{p-1}) \in X$.

L'équation (3.10) donne par récurrence pour $k \in \mathbb{N}$
$$\sigma^k \cdot (x_1, \dots, x_p) \in X \quad \forall\, (x_1, \dots, x_p) \in X\,.$$

Ainsi, pour tous $(x_1, x_2, \dots, x_p) \in X$ et $\sigma^k \in \langle \sigma \rangle$, on a $\sigma^k \cdot (x_1, \dots, x_p) \in X$. L'action (3.9) définit donc une application
$$\langle \sigma \rangle \times X \longrightarrow X. \tag{3.11}$$

L'application (3.11) est une action de $\langle \sigma \rangle$ sur $X$.

On considère cette action de $\langle \sigma \rangle$ sur $X$. D'après le Théorème 3.1, le cardinal d'une orbite de $X$ divise l'ordre du groupe agissant, qui est ici $p$. Donc le cardinal d'une orbite de $X$ est 1 ou $p$ puisque $p$ est premier. Une orbite est de cardinal 1 si

et seulement si c'est $\{(x_1, x_2, \dots, x_p)\}$ avec $x_1 = x_2 = \dots = x_p = x$, $x^p = e$ *i.e.* $\{(x, x, \dots, x)\}$ avec $x^p = e$.

$\{(e, e, \dots, e)\}$ est une orbite de cardinal 1. S'il n'existait pas d'autre orbite de cardinal 1, comme $X$ est réunion disjointe des orbites, on aurait $|X| = 1 + Np$ , *i.e.* $n^{p-1} = 1 + Np$ , où $N$ est le nombre des orbites différentes de $\{(e, e, \dots, e)\}$. Comme $p$ divise $n$, cela impliquerait que $p$ divise 1, ce qui est absurde, puisque $p \neq 1$. Donc il existe au moins une orbite de cardinal 1, différente de $\{(e, e, \dots, e)\}$, *i.e.* il existe $x \neq e$ tel que $x^p = e$. □

## 3.3 Exercices.

**Exercice 3.1.**

Montrer que l'on définit une action de $G = \mathcal{S}_n$ sur $X = \mathbb{R}[X_1, \dots, X_n]$ en posant :

$$(\sigma, f) \mapsto \sigma \cdot f \quad \forall\, \sigma \in G,\ \forall\, f \in X,$$

où

$$(\sigma \cdot f)(X_1, \dots, X_n) = f(X_{\sigma(1)}, \dots, X_{\sigma(n)}) \quad \forall\, \sigma \in G,\ \forall\, f \in X.$$

**Exercice 3.2.**

Soit $K$ un corps commutatif et $E = K_2[X]$ l'espace vectoriel des polynômes de degré $\leqslant 2$ en $X$ à coefficients dans $K$.

On pose pour $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL(2, K)$ et $P \in E$

$$(A \cdot P)(X) = (bX + d)^2\, P\left(\frac{aX + c}{bX + d}\right). \tag{3.12}$$

(i) Montrer que l'application $(A, P) \mapsto A \cdot P$ est une action de $GL(2, K)$ sur $E$.

(ii) Montrer que pour tout $A \in GL(2, K)$ l'application $f_A : E \to E$ définie par $f_A(P) = A \cdot P$ est un endomorphisme de $E$ et écrire sa matrice dans la base canonique $\mathcal{B} = (1, X, X^2)$ de $E$.

(iii) Montrer que pour tout $\lambda \in K, \lambda \neq 0$ et tout $P \in E$

$$(\lambda\, A) \cdot P = \lambda^2\, (A \cdot P) \tag{3.13}$$

et en déduire que

$$\left(\frac{1}{\lambda}\, A\right) \cdot (\lambda^2\, P) = A \cdot P \tag{3.14}$$

(iv) Pour $P = x + yX + zX^2 \in E, P \neq 0$, notons $\Delta_P = y^2 - 4xz$ le discriminant de $P$. Montrer que

$$\Delta_{A \cdot P} = (\det A)^2\, \Delta_P \quad \forall P \in E. \tag{3.15}$$

(v) On prend $K = \mathbb{C}$. Déterminer les orbites de $E$ sous $GL(2, \mathbb{C})$.

(vi) On prend $K = \mathbb{R}$. Déterminer les orbites de $E$ sous $GL(2, \mathbb{R})$ et les représenter graphiquement.

(vii) Reprendre les questions (v) et (vi) avec $SL(2, K)$ ($K = \mathbb{C}$ ou $K = \mathbb{R}$) pour l'action restreinte déduite de celle de $GL(2, K)$.

**Indication.**

(i) Soit $P = x + yX + zX^2 \in E$ et $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL(2, K)$. Alors

$$\begin{aligned}(bX+d)^2 P\left(\frac{aX+c}{bX+d}\right) &= (bX+d)^2 \left(x + y\,\frac{aX+c}{bX+d} + z\left(\frac{aX+c}{bX+d}\right)^2\right) \\ &= x(bX+d)^2 + y(aX+c)(bX+d) + z(aX+c)^2\end{aligned}$$

donc $A \cdot P \in E \ \forall P \in E$. Si $I$ désigne la matrice identité, on a immédiatement $I \cdot P = P \ \forall P \in E$. Enfin pour tous $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, A' = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \in GL(2, K)$ et $P \in E$, on a :

$$\begin{aligned}(A \cdot (A' \cdot P))(X) &= (bX+d)^2 (A' \cdot P)\left(\frac{aX+c}{bX+d}\right) \\ &= (bX+d)^2 \left(b'\,\frac{aX+c}{bX+d} + d'\right)^2 P\left(\frac{a'\,\frac{aX+c}{bX+d} + c'}{b'\,\frac{aX+c}{bX+d} + d'}\right) \\ &= ((ab' + bd')X + cb' + dd')^2 P\left(\frac{(aa' + bc')X + ca' + dc'}{(ab' + bd')X + cb' + dd'}\right) \\ &= ((AA') \cdot P)(X).\end{aligned}$$

Donc on a bien une action de $GL(2, K)$ sur $E$.

(ii) La linéarité de $f_A$ est immmédiate sur la formule (3.12) : pour $\lambda, \mu \in K$ et $P, Q \in E$,

$$\begin{aligned}(A \cdot (\lambda P + \mu Q))(X) &= (bX+d)^2 (\lambda P + \mu Q)\left(\frac{aX+c}{bX+d}\right) \\ &= \lambda(bX+d)^2 P\left(\frac{aX+c}{bX+d}\right) + \mu(bX+d)^2 Q\left(\frac{aX+c}{bX+d}\right) \\ &= \lambda(A \cdot P)(X) + \mu(A \cdot Q)(X),\end{aligned}$$

*i.e.* $f_A(\lambda P + \mu Q) = \lambda f_A(P) + \mu f_A(Q)$.

On a pour $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL(2, K)$ en notant $e_1 = 1, e_2 = X, e_3 = X^2$ :

$$A \cdot e_1 = (bX+d)^2 \tag{3.16}$$

$$A \cdot e_2 = (bX+d)(aX+c) \tag{3.17}$$

$$A \cdot e_3 = (aX+c)^2. \tag{3.18}$$

La matrice de $f_A$ dans la base $\mathcal{B}$ est donc

$$\mathcal{M}(f_A, \mathcal{B}) = \begin{pmatrix} d^2 & cd & c^2 \\ 2bd & ad+bc & 2ac \\ b^2 & ab & a^2 \end{pmatrix}. \tag{3.19}$$

(iii) Pour $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL(2, K)$ et $\lambda \in K, \lambda \neq 0$, on a

$$\det(\lambda A) = \lambda^2 \det A \neq 0$$

donc $\lambda A \in GL(2,K)$. La formule (3.12) donne

$$\begin{aligned}((\lambda A)\cdot P)(X) &= (\lambda bX+\lambda d)^2\, P\left(\frac{\lambda aX+\lambda c}{\lambda bX+\lambda d}\right)\\ &= \lambda^2(bX+d)^2\, P\left(\frac{aX+c}{bX+d}\right)\\ &= \lambda^2(A\cdot P)(X).\end{aligned}$$

C'est l'équation (3.13). Maintenant,

$$\left(\frac{1}{\lambda}A\right)\cdot(\lambda^2 P)=\frac{1}{\lambda^2}\left(A\cdot(\lambda^2 P)\right)=A\cdot P$$

par linéarité de l'application $f_A$. C'est l'équation (3.14).

(iv) Si $P$ a pour composantes dans la base canonique $x,y,z$, les composantes de $A\cdot P$ sont, d'après (3.19), $x',y',z'$ avec

$$\begin{aligned}x' &= d^2x+cdy+c^2z\\ y' &= 2bdx+(ad+bc)y+2acz\\ z' &= b^2x+aby+a^2z.\end{aligned}$$

On a alors

$$\begin{aligned}\Delta_{A\cdot P} &= (y')^2-4x'z'\\ &= y^2((ad+bc)^2-4abcd)-4xz(a^2d^2+b^2c^2-2abcd)\\ &= (y^2-4xz)(ad-bc)^2\\ &= (\det A)^2\,\Delta_P.\end{aligned}$$

(v) L'orbite du polynôme 0 est $\{0\}$. Soit $P=x+yX+zX^2\in\mathbb{C}_2[X], P\neq 0$. Deux cas sont possibles sur $\mathbb{C}$ :

• $\Delta_P=0$. Dans ce cas, $z=0$ impliquerait $y=0$. Donc $P$ est constant ou de degré 2 avec une racine double.

Si $P$ est la constante $x\in\mathbb{C}^*$, on a $P=A\cdot e_1$ d'après la formule (3.16), avec $A=\begin{pmatrix}a & 0\\ c & d\end{pmatrix}\in GL(2,\mathbb{C})$ pour tous $a,c,d\in\mathbb{C}$ tels que $a\neq 0$ et $d^2=x$ (on peut en particulier prendre $a=d^{-1}$ qui donne $\det A=1$). On a donc $P\in\mathcal{O}(e_1)$.

Si $P$ est de degré 2 avec une racine double, il s'écrit $P=z(X-\alpha)^2$ avec $\alpha\in\mathbb{C}$ , $z\in\mathbb{C}^*$. Si $b\in\mathbb{C}^*$ est tel que $b^2=z$, on a $P=A\cdot e_1$ avec $A=\begin{pmatrix}a & b\\ c & -\alpha b\end{pmatrix}\in GL(2,\mathbb{C})$ pour tous $a,c\in\mathbb{C}$ tels que $a\alpha+c\neq 0$ (on peut en particulier prendre $a=0$ et $c=-b^{-1}$ qui donne $\det A=1$). On a donc encore $P\in\mathcal{O}(e_1)$.

On a donc obtenu que, pour $P\neq 0$, la condition $\Delta_P=0$ implique $P\in\mathcal{O}(e_1)$. Or d'après la formule (3.16) tout point de l'orbite de $e_1$ vérifie $P\neq 0$ et $\Delta_P=0$. Donc l'orbite $\mathcal{O}(e_1)$ est l'ensemble des polynômes $P\neq 0$ qui vérifient $\Delta_P=0$. (Elle contient en particulier $e_3$, donc est identique à l'orbite de $e_3$).

• $\Delta_P\neq 0$. Dans ce cas $z=0$ impliquerait $y\neq 0$. Donc $P$ est de degré 1, ou de degré 2 avec deux racines distinctes.

Si $P$ est de degré 1, $P = x + yX, y \neq 0$, la matrice $A = \begin{pmatrix} 0 & y \\ 1 & x \end{pmatrix} \in GL(2, \mathbb{C})$ est telle que $P = A \cdot e_2$ d'après la formule (3.17). On a donc $P \in \mathcal{O}(e_2)$.

Si $P$ est de degré 2 avec deux racines distinctes, il s'écrit $P = z(X - \alpha)(X - \beta)$ avec $\alpha, \beta \in \mathbb{C}$ , $\alpha \neq \beta$ , $z \in \mathbb{C}^*$. Alors $P = A \cdot e_2$ avec en particulier la matrice $A = \begin{pmatrix} z & 1 \\ -z\beta & -\alpha \end{pmatrix} \in GL(2, \mathbb{C})$ dont le déterminant est $\det A = z(\beta - \alpha)$. On a donc dans ce cas aussi $P \in \mathcal{O}(e_2)$.

On a donc obtenu que la condition $\Delta_P \neq 0$ implique $P \in \mathcal{O}(e_2)$. Or d'après la formule (3.15), pour tout point $A \cdot e_2$ de l'orbite de $e_2$ on a $\Delta_{A \cdot e_2} = (\det A)^2 \neq 0$. Donc l'orbite $\mathcal{O}(e_2)$ est l'ensemble des polynômes qui vérifient $\Delta_P \neq 0$.

En résumé, il y a 3 orbites pour $GL(2, \mathbb{C})$ :

▶ l'orbite triviale $\{0\}$;

▶ l'orbite $\mathcal{O}(e_1)$, ses éléments sont les polynômes non nuls qui vérifient $\Delta_P = 0$;

▶ l'orbite $\mathcal{O}(e_2)$, ses éléments sont les polynômes qui vérifient $\Delta_P \neq 0$.

(vi) Sur $K = \mathbb{R}$, il y a 3 cas possibles pour $P \neq 0$ :

• $\Delta_P = 0$. Dans ce cas, comme précédemment, $P$ est constant ou de degré 2 avec une racine double réelle.

Si $P$ est la constante $x \in \mathbb{R}^*$, on a $P = A \cdot (\varepsilon e_1)$ d'après la formule (3.16), avec $\varepsilon = \pm 1$ suivant le signe de $x$ et $A = \begin{pmatrix} a & 0 \\ c & \sqrt{|x|} \end{pmatrix} \in GL(2, \mathbb{R})$ pour tous $a, c \in \mathbb{R}$, $a \neq 0$ (on peut en particulier prendre $a = (\sqrt{|x|})^{-1}$ qui donne $\det A = 1$).

On a donc $P \in \mathcal{O}(e_1)$ ou $P \in \mathcal{O}(-e_1)$ suivant le signe de $x$.

Si $P$ est de degré 2 avec une racine double réelle, il s'écrit $P = z(X - \alpha)^2$ avec $\alpha \in \mathbb{R}, z \in \mathbb{R}^*$. Alors $P = A \cdot (\varepsilon e_1)$ avec $\varepsilon = \pm 1$ suivant le signe de $z$ et $A = \begin{pmatrix} a & \sqrt{|z|} \\ c & -\alpha\sqrt{|z|} \end{pmatrix} \in GL(2, \mathbb{R})$ pour tous $a, c \in \mathbb{R}$ tels que $a\alpha + c \neq 0$ (on peut en particulier prendre $a = 0$ et $c = -(\sqrt{|z|})^{-1}$ qui donne $\det A = 1$). Notons que le signe de $z$ est celui de $x + z$ puisque $4xz = y^2 \geqslant 0$. On a donc dans ce cas $P \in \mathcal{O}(e_1)$ ou $P \in \mathcal{O}(-e_1)$ suivant le signe de $x + z$.

Ainsi, pour $P \neq 0$, la condition $\Delta_P = 0$ implique $P \in \mathcal{O}(e_1)$ ou $P \in \mathcal{O}(-e_1)$ suivant le signe de $x + z$. Or comme précédemment, d'après la formule (3.16), pour tout point $A \cdot e_1$ de l'orbite de $e_1$ et tout point $-A \cdot e_1$ de l'orbite de $-e_1$ on a $A \cdot e_1 \neq 0$ et $\Delta_{A \cdot e_1} = \Delta_{-A \cdot e_1} = 0$. De plus si $P = A \cdot e_1 \in \mathcal{O}(e_1)$ (*resp.* $P = -A \cdot e_1 \in \mathcal{O}(-e_1)$), on a $P = x + yX + zX^2 = (bX + d)^2$ (*resp.* $P = x + yX + zX^2 = -(bX + d)^2$) avec $b, d \in \mathbb{R}$, donc $x + z > 0$ (*resp.* $x + z < 0$). L'orbite $\mathcal{O}(e_1)$ (*resp.* $\mathcal{O}(-e_1)$) est donc l'ensemble des $P \in E$ tels que $\Delta_P = 0$ et $x + z > 0$ (*resp.* $x + z < 0$). L'ensemble des polynômes $P \neq 0$ qui vérifient $\Delta_P = 0$ est ainsi la réunion des deux orbites distinctes $\mathcal{O}(e_1)$ et $\mathcal{O}(-e_1)$.

• $\Delta_P > 0$. Dans ce cas, $P$ est de degré 1, ou de degré 2 avec deux racines réelles distinctes. La situation est exactement la même que dans le cas complexe.

Si $P$ est de degré 1, $P = x + yX, y \neq 0$, la matrice $A = \begin{pmatrix} 0 & y \\ 1 & x \end{pmatrix} \in GL(2, \mathbb{R})$ est telle que $P = A \cdot e_2$ d'après la formule (3.17). On a donc $P \in \mathcal{O}(e_2)$.

Si $P$ est de degré deux avec deux racines réelles distinctes, il s'écrit $P = z(X - \alpha)(X - \beta)$ avec $\alpha, \beta \in \mathbb{R}, \alpha \neq \beta, z \in \mathbb{R}^*$. Alors $P = A \cdot e_2$ avec

$A = \begin{pmatrix} z & 1 \\ -z\beta & -\alpha \end{pmatrix} \in GL(2, \mathbb{R})$ dont le déterminant est $\det A = z(\beta - \alpha)$. On a donc dans ce cas aussi $P \in \mathcal{O}(e_2)$.

Si $\Delta_P > 0$, on a donc toujours $P \in \mathcal{O}(e_2)$. Or d'après la formule (3.15), pour tout point $A \cdot e_2$ de l'orbite de $e_2$ on a $\Delta_{A \cdot e_2} = (\det A)^2 > 0$. Donc l'orbite $\mathcal{O}(e_2)$ est l'ensemble des polynômes qui vérifient $\Delta_P > 0$.

• $\Delta_P < 0$. Dans ce cas, $P$ est de degré 2, avec deux racines complexes conjuguées distinctes $\alpha \pm i\beta$, $\alpha, \beta \in \mathbb{R}$, $\beta \neq 0$. Il s'écrit

$$\begin{aligned} P &= z(X - \alpha - i\beta)(X - \alpha + i\beta) \\ &= z(X^2 - 2\alpha X + \alpha^2 + \beta^2) \\ &= z((X - \alpha)^2 + \beta^2) \quad (z \in \mathbb{R}^*). \end{aligned}$$

Or d'après les formules (3.16) et (3.18),

$$A(e_1 + e_3) = (bX + d)^2 + (aX + c)^2$$

pour $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL(2, \mathbb{R})$. Donc $P = A \cdot (\varepsilon(e_1 + e_3))$ avec $\varepsilon = \pm 1$ suivant le signe de $z$ et $A = \sqrt{|z|} \begin{pmatrix} 0 & 1 \\ \beta & -\alpha \end{pmatrix} \in GL(2, \mathbb{R})$. Notons que le signe de $z$ est celui de $x + z$ puisque $4xz > y^2 \geqslant 0$. On a donc $P \in \mathcal{O}(e_1 + e_3)$ ou $P \in \mathcal{O}(-(e_1 + e_3))$ suivant le signe de $x + z$. Or comme précédemment, d'après la formule (3.15), tout point $P$ de l'orbite de $e_1 + e_3$ ou de l'orbite de $-(e_1 + e_3)$ vérifie $\Delta_P < 0$. De plus si $P \in \mathcal{O}(e_1 + e_3)$ (*resp.* $P \in \mathcal{O}(-(e_1 + e_3))$), on a $P = x + yX + zX^2 = (bX + d)^2 + (aX + c)^2$ (*resp.* $P = x + yX + zX^2 = -(bX + d)^2 - (aX + c)^2$) avec $a, b, c, d \in \mathbb{R}$, donc $x + z > 0$ (*resp.* $x + z < 0$). L'orbite $\mathcal{O}(e_1 + e_3)$ (*resp.* $\mathcal{O}(-(e_1 + e_3))$) est donc l'ensemble des $P \in E$ tels que $\Delta_P < 0$ et $x + z > 0$ (*resp.* $x + z < 0$). L'ensemble des polynômes qui vérifient $\Delta_P < 0$ est ainsi la réunion des deux orbites distinctes $\mathcal{O}(e_1 + e_3)$ et $\mathcal{O}(-(e_1 + e_3))$.

En résumé, il y a 6 orbites pour $GL(2, \mathbb{R})$ :

▶ l'orbite triviale $\{0\}$ ;

▶ l'orbite $\mathcal{O}(e_1)$, ses éléments sont les polynômes non nuls qui vérifient $\Delta_P = 0$ et $x + z > 0$;

▶ l'orbite $\mathcal{O}(-e_1)$, ses éléments sont les polynômes non nuls qui vérifient $\Delta_P = 0$ et $x + z < 0$ ;

▶ l'orbite $\mathcal{O}(e_2)$, ses éléments sont les polynômes qui vérifient $\Delta_P > 0$ ;

▶ l'orbite $\mathcal{O}(e_1 + e_3)$, ses éléments sont les polynômes non nuls qui vérifient $\Delta_P < 0$ et $x + z > 0$ ;

▶ l'orbite $\mathcal{O}(-(e_1 + e_3))$, ses éléments sont les polynômes non nuls qui vérifient $\Delta_P < 0$ et $x + z < 0$.

*Représentation géométrique des orbites.* Munissons $E$ du produit scalaire pour lequel la base $(\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3})$ est orthonormée. La forme quadratique $\Delta_P = y^2 - 4xz$ s'écrit

$$\Delta_P = (x - z)^2 + y^2 - (x + z)^2 = 2\left(\frac{x - z}{\sqrt{2}}\right)^2 + y^2 - 2\left(\frac{x + z}{\sqrt{2}}\right)^2 = 2\xi^2 + \eta^2 - 2\zeta^2,$$

$(\xi, \eta, \zeta)$ désignant les coordonnées du polynôme $P$ dans la base orthonormée $(\mathbf{f_1}, \mathbf{f_2}, \mathbf{f_3})$ définie par

$$\mathbf{f_1} = \frac{\mathbf{e_1} - \mathbf{e_3}}{\sqrt{2}}, \ \mathbf{f_2} = \mathbf{e_2}, \ \mathbf{f_3} = \frac{\mathbf{e_1} + \mathbf{e_3}}{\sqrt{2}}.$$

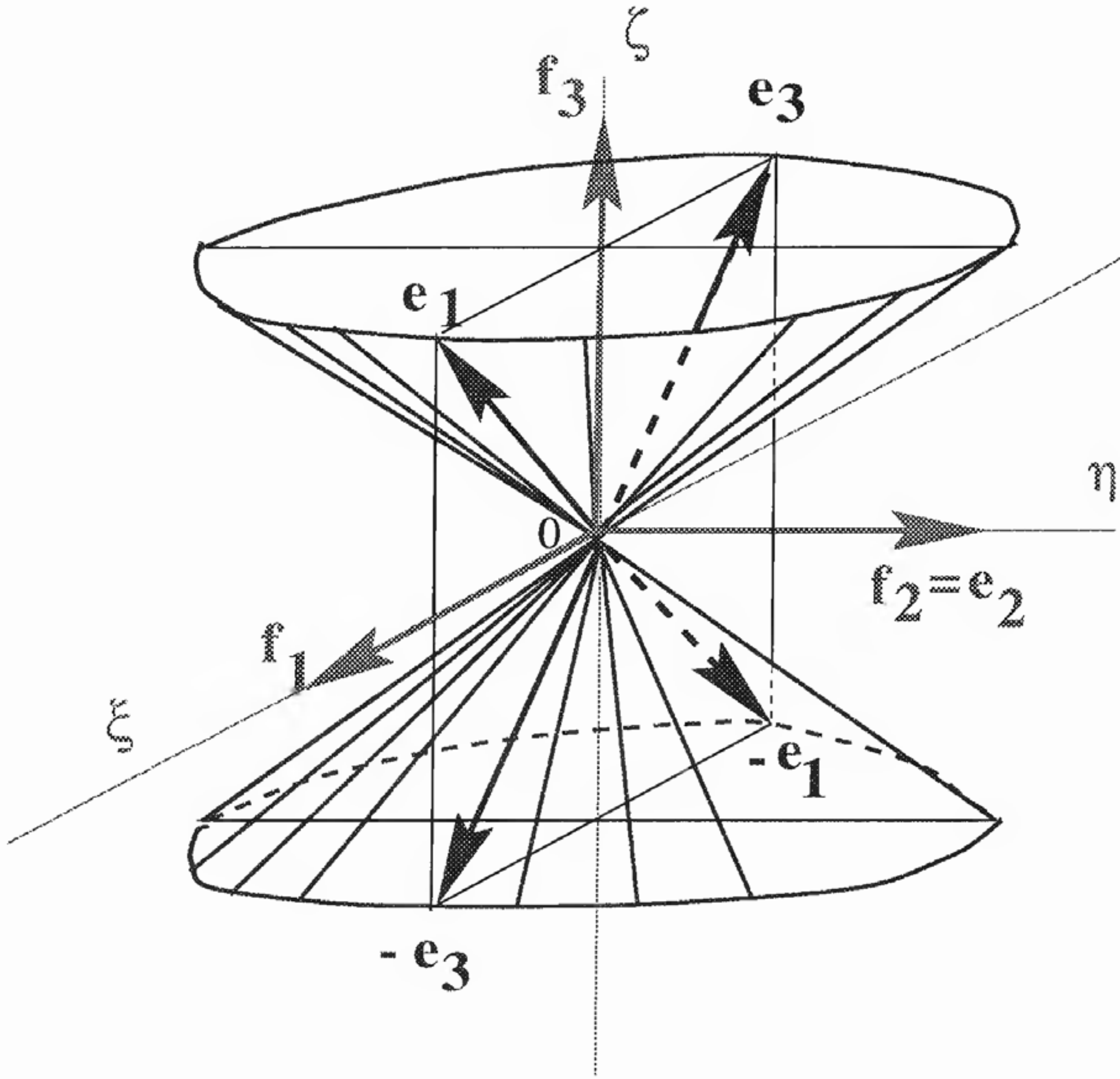



FIG. 3.1: *Les orbites de l'action de $GL(2,\mathbb{R})$ sur $\mathbb{R}_2[X]$.*

En effet, le changement de bases est en fait un changement de bases orthonormées dans le plan des $x, z$, donc $\eta = y$, et la matrice de passage $P$ de $(\mathbf{e_1}, \mathbf{e_3})$ à $(\mathbf{f_1}, \mathbf{f_3})$ est $P = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \in SO(2)$ donc $\begin{pmatrix} \xi \\ \zeta \end{pmatrix} = P^{-1}\begin{pmatrix} x \\ z \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} x-z \\ x+z \end{pmatrix}$.

La surface d'équation $\Delta_P = 0$ a pour équation dans la base $(\mathbf{f_1}, \mathbf{f_2}, \mathbf{f_3})$ :

$$2\xi^2 + \eta^2 - 2\zeta^2 = 0. \tag{3.20}$$

C'est un cône d'axe $\mathbb{R}\,\mathbf{f_3}$. Pour $\zeta_0 \neq 0$ fixé, on obtient $\frac{\xi^2}{\zeta_0{}^2} + \frac{\eta^2}{2\zeta_0{}^2} = 1$. La section du cône par le plan $\zeta = \zeta_0$ est donc une ellipse dont le grand axe est $\sqrt{2}|\zeta_0|$ suivant $\eta$ et le petit axe est $|\zeta_0|$ suivant $\xi$.

Les orbites sont ainsi représentées comme (voir Fig. 3.1) :

- l'orbite triviale $\{0\}$ ;
- la nappe supérieure ($\zeta > 0$) du cône (orbite $\mathcal{O}(e_1)$) ;
- la nappe inférieure ($\zeta < 0$) du cône (orbite $\mathcal{O}(-e_1)$) ;
- "l'extérieur" du cône (orbite $\mathcal{O}(e_2)$) ;
- "l'intérieur" ($\zeta > 0$) du cône supérieur (orbite $\mathcal{O}(e_1 + e_3)$) ;

► "l'intérieur" ($\zeta < 0$) du cône inférieur (orbite $\mathcal{O}(-(e_1+e_3))$).

(vii) Si $A \in SL(2,K)$, on a d'après (3.15) $\Delta_{A \cdot P} = \Delta_P \ \forall P \in E$. Le discriminant est donc constant sur chaque orbite.

Prenons d'abord $K = \mathbb{C}$. On va voir que les orbites pour $SL(2,\mathbb{C})$ sont alors :

► l'orbite triviale $\{0\}$;

► l'orbite $\mathcal{O}(e_1)$, ses éléments sont les polynômes non nuls qui vérifient $\Delta_P = 0$;

► l'orbite $\mathcal{O}(\lambda\, e_2)$, pour tout $\lambda \in \mathbb{C}^*$, tel que $0 \leqslant \operatorname{Arg} \lambda < \pi$, où $\operatorname{Arg} \lambda$ désigne la détermination principale de l'argument de $\lambda$ (voir Chap. 7 , section 7.10). Ses éléments sont les polynômes qui vérifient $\Delta_P = \lambda^2$.

Soit $P \in \mathbb{C}_2[X], P \neq 0$. Si $\Delta_P = 0$, on a vu au cours de la démonstration du (v) qu'il existe une matrice $A \in SL(2,\mathbb{C})$ telle que $P = A \cdot e_1$. Supposons maintenant $\Delta_P \neq 0$. Il existe un unique $\lambda \in \mathbb{C}$ tel que $0 \leqslant \operatorname{Arg} \lambda < \pi$ et $\Delta_P = \lambda^2$. On sait d'après (v) qu'il existe $A \in GL(2,\mathbb{C})$ tel que $P = A \cdot e_2$. Alors $\Delta_P = (\det A)^2$, donc $\det A = \pm\lambda$. Si $\det A = -\lambda$, on introduit la matrice

$$C = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{3.21}$$

qui est telle que $\det C = -1$ et $C \cdot e_2 = C \cdot X = X^2 \left(\frac{1}{X}\right) = X = e_2$. On aura $P = A \cdot e_2 = A \cdot (C \cdot e_2) = (AC) \cdot e_2$. Or la matrice $A' = AC$ vérifie $\det A' = \lambda$. En remplaçant $A$ par $A'$, on peut donc supposer que $\det A = \lambda$. Soit $\mu \in \mathbb{C}$ tel que $\mu^2 = \lambda$. Alors $B = \frac{1}{\mu} A \in SL(2,\mathbb{C})$. D'après (3.14), $P = A \cdot e_2 = B \cdot (\mu^2 e_2) = B \cdot (\lambda\, e_2)$. Cela prouve que $P \in \mathcal{O}(\lambda\, e_2)$. D'où la classification des orbites.

Prenons maintenant $K = \mathbb{R}$. On va voir que les orbites pour $SL(2,\mathbb{R})$ sont (voir Fig. 3.2) :

► l'orbite triviale $\{0\}$;

► la nappe supérieure ($\zeta > 0$) du cône (orbite $\mathcal{O}(e_1)$) ;

► la nappe inférieure ($\zeta < 0$) du cône (orbite $\mathcal{O}(-e_1)$) ;

► pour tout $\lambda > 0$, l'hyperboloïde à 1 nappe $\Delta_P = y^2 - 4xz = \lambda^2$ (orbite $\mathcal{O}(\lambda\, e_2)$) ;

► pour tout $\lambda > 0$, la nappe supérieure ($\zeta > 0$) de l'hyperboloïde à 2 nappes $\Delta_P = y^2 - 4xz = -4\lambda^2$ (orbite $\mathcal{O}(\lambda(e_1+e_3))$) ;

► pour tout $\lambda > 0$, la nappe inférieure ($\zeta < 0$) de l'hyperboloïde à 2 nappes $\Delta_P = y^2 - 4xz = -4\lambda^2$ (orbite $\mathcal{O}(-\lambda(e_1+e_3))$).

Soit $P \in \mathbb{R}_2[X], P \neq 0$. Si $\Delta_P = 0$, on a vu au cours de la démonstration du (vi) qu'il existe une matrice $A \in SL(2,\mathbb{R})$ telle que $P = A \cdot e_1$ ou $P = -A \cdot e_1$ suivant le signe de $x + z$.

Supposons $\Delta_P = \lambda^2 > 0$ ($\lambda > 0$). On sait d'après (vi) qu'il existe $A \in GL(2,\mathbb{R})$ tel que $P = A \cdot e_2$. Alors $\Delta_P = (\det A)^2$, donc $\det A = \pm\lambda$. Le cas $\det A = -\lambda$, se ramène au cas $\det A = \lambda$ comme dans le cas complexe en introduisant la matrice (3.21). Supposons donc $\det A = \lambda$. Alors $B = \frac{1}{\sqrt{\lambda}} A \in SL(2,\mathbb{R})$. D'après (3.14), $P = A \cdot e_2 = B \cdot (\lambda\, e_2)$. Cela prouve que $P \in \mathcal{O}(\lambda\, e_2)$. On en déduit que l'orbite de $\lambda\, e_2$ est l'hyperboloïde à 1 nappe $\Delta_P = y^2 - 4xz = \lambda^2$.

Supposons $\Delta_P = -4\lambda^2 < 0$ ($\lambda > 0$). On sait d'après (vi) qu'il existe $A \in GL(2,\mathbb{R})$ tel que $P = A \cdot (\varepsilon(e_1 + e_3))$ avec $\varepsilon = \pm 1$ suivant le signe de $x + z$. Alors $\Delta_P = -4(\det A)^2$, donc $\det A = \pm\lambda$. Si $\det A = -\lambda$, on introduit la matrice

$$S = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

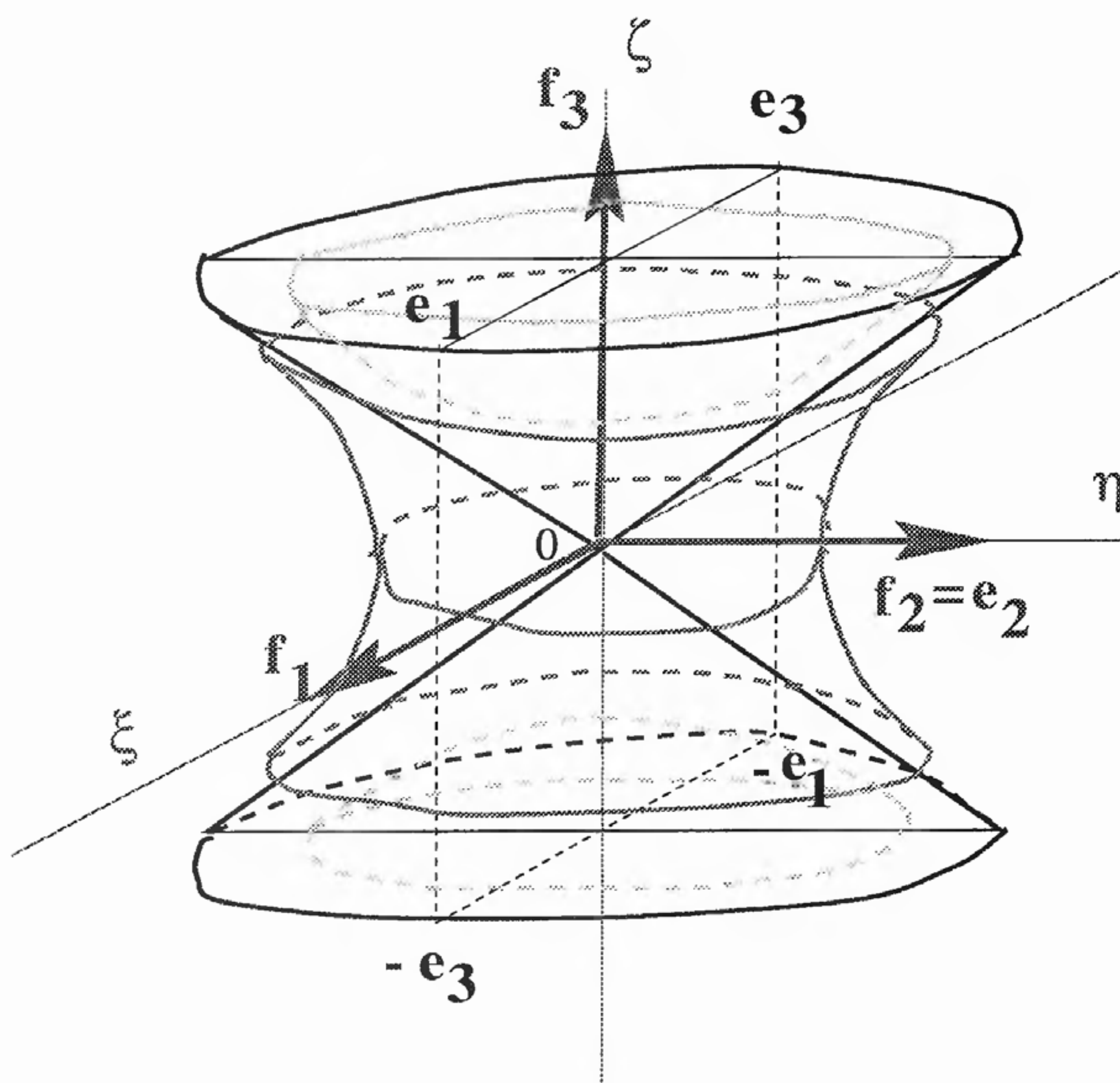



FIG. 3.2: *Les orbites de l'action de* $SL(2,\mathbb{R})$ *sur* $\mathbb{R}_2[X]$.

qui est telle que $\det S = -1$ et

$$S \cdot (e_1 + e_3) = \frac{1}{2}\left((X-1)^2 + (X+1)^2\right) = 1 + X^2 = e_1 + e_3.$$

On aura $P = A \cdot (\varepsilon(e_1 + e_3)) = A \cdot \varepsilon\left(S \cdot (e_1 + e_3)\right) = (AS) \cdot (\varepsilon(e_1 + e_3))$. La matrice $A' = AS$ vérifie alors $\det A' = \lambda$. On peut donc supposer que $\det A = \lambda$. Alors $B = \frac{1}{\sqrt{\lambda}} A \in SL(2, \mathbb{R})$. D'après (3.14), $P = A \cdot (\varepsilon(e_1 + e_3)) = B \cdot (\lambda\varepsilon(e_1 + e_3))$. Cela prouve que $P \in \mathcal{O}(\lambda\varepsilon(e_1 + e_3))$. On en déduit que l'orbite de $\lambda(e_1 + e_3)$ (*resp.* $-\lambda(e_1 + e_3)$) est la nappe supérieure (resp. inférieure) de l'hyperboloïde à 2 nappes $\Delta_P = y^2 - 4xz = -4\lambda^2$. D'où la classification des orbites.

**Exercice 3.3.**

Soit $p$ un nombre premier et $G$ un groupe fini d'ordre $p^s$, $s \in \mathbb{N}^*$. Montrer que le centre de $G$ n'est pas réduit à l'élément neutre.

**Indication.**

L'équation des classes s'écrit

$$|G| = |Z(G)| + \sum_{i=1}^{k} |\Omega_i|.$$

où $Z(G)$ est le centre de $G$ et les $\Omega_i$ sont les classes de conjugaison non réduites à un élément. D'après le Th. 3.1, $|\Omega_i|$ divise $|G| = p^s$. Comme $p$ est premier, on a donc $|\Omega_i| = p^{s_i}$ avec $s_i \in \mathbb{N}$, et $s_i \geqslant 1$ puisque la classe $\Omega_i$ n'est pas réduite à un élément. Si l'on avait $Z(G) = \{e\}$, on aurait donc

$$p^s = 1 + \sum_{i=1}^{k} p^{s_i}$$

et $p$ diviserait 1 ce qui est faux. Donc $Z(G) \neq \{e\}$.

**Exercice 3.4.**

Soit $G$ un groupe, et $Z(G)$ son centre.

(i) Montrer que si le groupe quotient $G/Z(G)$ est monogène, alors $G$ est abélien.

(ii) Dans toute la suite, on suppose que $G$ est fini de cardinal $|G| = p^2$, avec $p$ un nombre premier.

Montrer que $G$ est abélien en utilisant l'exercice 3.3 et la question précédente.

(iii) On suppose que $G$ n'est pas cyclique. Soit $x \in G$ tel que $x$ soit différent de l'élément neutre $e$.

(a) Quel est l'ordre de $x$?

(b) Soit $y \in G$ tel que $y \notin \langle x \rangle$. Quel est l'ordre de $y$? Montrer que $\langle x \rangle \cap \langle y \rangle = \{e\}$.

(c) Montrer que les éléments

$$x^i y^j,\ 0 \leqslant i \leqslant p-1,\ 0 \leqslant j \leqslant p-1$$

sont deux-à-deux distincts.

(d) En déduire que $G$ est isomorphe au produit direct $\mathbb{Z}_p \times \mathbb{Z}_p$

(iv) Conclure qu'il n'y a, à isomorphisme près, que 2 groupes d'ordre $p^2$ ($p$ premier), à savoir à $\mathbb{Z}_{p^2}$ et $\mathbb{Z}_p \times \mathbb{Z}_p$.

**Indication.**

(i) Par hypothèse, il existe un élément $a \in G$ tel que $G/Z(G) = \langle [a] \rangle$ , où $[a] = a\,Z(G)$ est la classe à gauche de $a$ modulo $Z(G)$. Soient $x, y \in G$. Il existe

$m, n \in \mathbb{N}$ tels que $[x] = [a]^m$ et $[y] = [a]^n$. Or $[a]^m = [a^m]$ et $[a]^n = [a^n]$. Donc $[x] = [a^m]$ et $[y] = [a^n]$, et par conséquent il existe $h, k \in Z(G)$ tels que $x = a^m h$ et $y = a^n k$. Alors $xy = a^m h a^n k = a^{m+n} hk = a^{m+n} kh = a^n k a^m h = yx$. Comme $x, y \in G$ sont arbitraires, $G$ est donc abélien.
(ii) D'après l'exercice 3.3, $Z(G) \neq \{e\}$. Or d'après le Théorème de Lagrange, $|Z(G)|$ est un diviseur de $p^2$. Donc $|Z(G)| = p$ ou $p^2$. Si $|Z(G)| = p^2$, $Z(G) = G$ et $G$ est donc abélien. Si $|Z(G)| = p$, $|G/Z(G)| = |G|/|Z(G)| = p$. Donc $G/Z(G) \cong \mathbb{Z}_p$ (Voir ex. 1.4) et $G/Z(G)$ est cyclique. D'après la question (i), G est alors abélien. Ainsi $G$ est abélien dans les deux cas.
(iii)(a) L'ordre de $x$ est un diviseur de $p^2$ différent de 1 puisque $x \neq e$, et différent de $p^2$, puisque $G \neq \langle x \rangle$ du fait que $G$ n'est pas cyclique. Donc l'ordre de $x$ est $p$.
(b) On a $y \neq e$ puisque $y \notin \langle x \rangle$. Donc par le même raisonnement que pour $x$, l'ordre de $y$ est $p$. D'autre part, $\langle x \rangle \cap \langle y \rangle$ est un sous groupe de $\langle y \rangle$ donc son ordre divise $p$. Si l'ordre était $p$ on aurait $\langle x \rangle \cap \langle y \rangle = \langle y \rangle$ ce qui contredirait l'hypothèse $y \notin \langle x \rangle$. Donc l'ordre est 1, *i.e.* $\langle x \rangle \cap \langle y \rangle = \{e\}$.
(c) Supposons $x^i y^j = x^{i'} y^{j'}$ avec $0 \leqslant i, i' \leqslant p-1$, $0 \leqslant j, j' \leqslant p-1$. Alors $x^{i-i'} = y^{j'-j}$ est un élément de $\langle x \rangle \cap \langle y \rangle = \{e\}$. Donc $x^{i-i'} = y^{j'-j} = e$. Comme les ordres de $x$ et $y$ sont tous les deux égaux à $p$, cela implique d'après le Th. 1.11 que $i - i' \in p\mathbb{Z}$ et $j' - j \in p\mathbb{Z}$. D'où $i = i'$ et $j = j'$.
(d) Considérons l'application $f : \langle x \rangle \times \langle y \rangle \to G$ définie par $f((z, w)) = zw \; \forall z \in \langle x \rangle, \; w \in \langle y \rangle$. Comme $G$ est abélien, $f$ est un homomorphisme de groupes : pour tous $z, z' \in \langle x \rangle$ , $w, w' \in \langle y \rangle$,

$$f((z, w)(z', w')) = f((zz', ww')) = zz'ww' = zwz'w' = f((z, w))f((z', w')).$$

D'après (c), les $p^2$ éléments

$$f((x^i, y^j)) = x^i y^j, \; 0 \leqslant i, i' \leqslant p-1, \; 0 \leqslant j, j' \leqslant p-1$$

dont deux-à-deux distincts. Comme $|G| = p^2$, $f$ est donc surjective. Comme $|\langle x \rangle \times \langle y \rangle| = |G|$, cela implique que $f$ est bijective. Ainsi $f$ est un isomorphisme de $\langle x \rangle \times \langle y \rangle$ sur $G$. Or $\langle x \rangle \cong \mathbb{Z}_p$ et $\langle y \rangle \cong \mathbb{Z}_p$. Donc $\langle x \rangle \times \langle y \rangle \cong \mathbb{Z}_p \times \mathbb{Z}_p$. On obtient ainsi $G \cong \mathbb{Z}_p \times \mathbb{Z}_p$.
(iv) Si $G$ est cyclique, on sait que $G \cong \mathbb{Z}_{p^2}$. Si $G$ n'est pas cyclique, on a vu à la question précédente que $G \cong \mathbb{Z}_p \times \mathbb{Z}_p$. Ainsi, un groupe $G$ d'ordre $p^2$ est isomorphe à $\mathbb{Z}_{p^2}$ ou $\mathbb{Z}_p \times \mathbb{Z}_p$. Il reste à montrer que $\mathbb{Z}_{p^2}$ et $\mathbb{Z}_p \times \mathbb{Z}_p$ ne sont pas isomorphes. Mais c'est immédiat car il n'y a pas d'élément d'ordre $p^2$ dans $\mathbb{Z}_p \times \mathbb{Z}_p$. Soit en effet $([a], [b]) \in \mathbb{Z}_p \times \mathbb{Z}_p$, $([a], [b]) \neq ([0], [0])$. On a

$$p([a], [b]) = (p[a], p[b]) = ([pa], [pb]) = ([0], [0])$$

donc l'ordre $k$ de $([a], [b])$ est un diviseur de $p$. Comme $p$ est premier et $k > 1$ puisque $([a], [b]) \neq ([0], [0])$, on a $k = p$. Tout élément $([a], [b]) \neq ([0], [0])$ de $\mathbb{Z}_p \times \mathbb{Z}_p$ est donc d'ordre $p$.

# Chapitre 4

# Espace affine, barycentre.

Dans tout ce chapitre, le corps de base est un corps commutatif $K$.

## 4.1 Espace affine.

### 4.1.1 Structure affine sur un espace vectoriel.

Soit $E$ un espace vectoriel sur $K$. L'addition vectorielle définit une action

$$(\mathbf{t}, \mathbf{a}) \mapsto \mathbf{t} \cdot \mathbf{a} = \mathbf{a} + \mathbf{t} \tag{4.1}$$

du groupe additif de l'espace vectoriel $E$ sur l'ensemble $E$. On a en effet :

$$\begin{aligned} \mathbf{a} + 0 &= \mathbf{a} \quad \forall \mathbf{a} \in E \\ (\mathbf{a} + \mathbf{t}) + \mathbf{s} &= \mathbf{a} + (\mathbf{t} + \mathbf{s}) \quad \forall \mathbf{a} \in E \quad \forall \mathbf{s}, \mathbf{t} \in E. \end{aligned}$$

Pour tous $\mathbf{a}, \mathbf{b} \in E$, $\mathbf{t} = \mathbf{b} - \mathbf{a}$ est l'unique élément de $E$ tel que $\mathbf{a} + \mathbf{t} = \mathbf{b}$.

### 4.1.2 Espace affine.

**Définition 4. 1.** *Soit $E$ un espace vectoriel sur $K$. Soit $X$ un ensemble dont les éléments seront appelés points et notés $A, B$ etc. On dit que $X$ est un espace affine associé à $E$, ou de direction $E$, si $X$ est muni d'une action notée*

$$(\mathbf{t}, M) \mapsto M + \mathbf{t} \tag{4.2}$$

*du groupe additif de l'espace vectoriel $E$ vérifiant la condition suivante.*

$$\textit{Pour tous } A, B \in X, \textit{ il existe un vecteur } \mathbf{t} \in E \textit{ unique tel que } B = A + \mathbf{t}. \tag{4.3}$$

*On notera ce vecteur $\mathbf{t} = \mathbf{AB}$ ou parfois $\mathbf{t} = B - A$. On appellera dimension de $X$ la dimension de $E$.*

**Exemple.** Soit $E$ un espace vectoriel sur $K$. L'action (4.1) du groupe additif de l'espace vectoriel $E$ munit l'ensemble $E$ d'une structure d'espace affine associé à l'espace vectoriel $E$. L'ensemble $E$ muni de cette structure d'espace affine est appelé *l'espace affine canoniquement associé à l'espace vectoriel $E$*, ou simplement *l'espace affine $E$*.

**Proposition 4. 1.** *Soit $X$ un espace affine associé à l'espace vectoriel $E$.*
*(i) Pour tout $A \in X$, l'application*

$$\varphi_A : E \longrightarrow X \tag{4.4}$$

*définie par $\varphi_A(\mathbf{t}) = A + \mathbf{t}$ pour tout $\mathbf{t} \in E$ est une bijection de $E$ sur $X$. L'image du vecteur nul $0$ par cette bijection est le point $A$.*
*(ii) La bijection réciproque $(\varphi_A)^{-1}$ de la bijection $\varphi_A : E \longrightarrow X$ est l'application*

$$\psi_A : X \longrightarrow E \tag{4.5}$$

*définie par $\psi_A(M) = \mathbf{AM} \quad \forall M \in X : \quad \psi_A \circ \varphi_A = Id_E \quad , \quad \varphi_A \circ \psi_A = Id_X$.*
*(iii) Pour tout $\mathbf{t} \in E$, on a $\varphi_A(\mathbf{s} + \mathbf{t}) = \varphi_A(\mathbf{s}) + \mathbf{t} \quad \forall \mathbf{s} \in E$, et $\psi_A(M + \mathbf{t}) = \psi_A(M) + \mathbf{t} \ \forall M \in X$.*
*(iv) Pour tous $A, B, C \in X$, on a : $\mathbf{AC} = \mathbf{AB} + \mathbf{BC}$ et $\mathbf{BA} = -\mathbf{AB}$.*

**Démonstration.**
(i) et (ii) sont une simple reformulation de (4.3).
(iii) $\varphi_A(\mathbf{s} + \mathbf{t}) = A + (\mathbf{s} + \mathbf{t}) = (A + \mathbf{s}) + \mathbf{t} = \varphi_A(\mathbf{s}) + \mathbf{t}$. On a alors $\varphi_A(\psi_A(M) + \mathbf{t}) = \varphi_A(\psi_A(M)) + \mathbf{t} = M + \mathbf{t}$ donc $\psi_A(M) + \mathbf{t} = \psi_A(M + \mathbf{t})$.
(iv) On a $A + (\mathbf{AB} + \mathbf{BC}) = (A + \mathbf{AB}) + \mathbf{BC} = B + \mathbf{BC} = C$ donc $\mathbf{AC} = \mathbf{AB} + \mathbf{BC}$. En prenant en particulier $C = A$, on obtient $0 = \mathbf{AB} + \mathbf{BA}$, *i.e.* $\mathbf{BA} = -\mathbf{AB}$. □

D'après la Prop. 4.1(i), si l'on fixe un point $A \in X$, la bijection $\psi_A : X \longrightarrow E$ permet d'identifier l'ensemble $X$ à l'ensemble $E$ en identifiant un point $M \in X$ au vecteur $\psi_A(M) = \mathbf{AM} \in E$. Alors d'après la Prop. 4.1(iii), pour tout $\mathbf{t} \in E$, le point $M + \mathbf{t}$ s'identifie au vecteur $\psi_A(M) + \mathbf{t} = \mathbf{AM} + \mathbf{t}$, *i.e.* l'action du groupe additif de l'espace vectoriel $E$ sur $X$ s'identifie à l'action (4.1). Avec cette identification, l'espace affine $X$ est ainsi l'espace affine canoniquement associé à $E$.

### 4.1.3 Repère affine.

**Définition 4. 2.** *Soit $X$ un espace affine associé à l'espace vectoriel $E$. On appelle repère affine de $X$ la donnée $(A, (\mathbf{e_i})_{1 \leqslant i \leqslant n})$ d'un point $A \in X$ appelé origine et d'une base $(\mathbf{e_i})_{1 \leqslant i \leqslant n}$ ($n = \dim E$) de l'espace vectoriel $E$.*

Lorsqu'un repère affine $(A, (\mathbf{e_i})_{1 \leqslant i \leqslant n})$ de $X$ est fixé, tout point $M \in X$ s'écrit de façon unique $M = A + \sum_{i=1}^n \lambda_i \mathbf{e_i}, \quad \lambda_i \in K$, d'après la Proposition 4.1(i). $\lambda_1, \cdots, \lambda_n$ (ou $(\lambda_1, \cdots, \lambda_n)$) sont appelées les coordonnées du point $M$.

L'espace affine $X$ est dit euclidien si l'espace vectoriel $E$ est euclidien. Dans ce cas, un repère affine $(A, (\mathbf{e_i})_{1 \leqslant i \leqslant n})$ est dit orthonormé si la base $(\mathbf{e_i})_{1 \leqslant i \leqslant n}$ de $E$ est orthonormée.

**Exemple.** Si $X$ désigne l'espace affine canoniquement associé à l'espace euclidien $E = \mathbb{R}^n$, soit $O$ le point de $X$ qui est le vecteur nul de $E$, et $(\mathbf{e_i})_{1 \leqslant i \leqslant n}$ la base orthonormée canonique de $E$. $(O, (\mathbf{e_i})_{1 \leqslant i \leqslant n})$ est un repère affine orthonormé. On dit que $O$ est l'origine canonique de $X$.

### 4.1.4 Application affine.

**Application affine dans un espace vectoriel.**

**Définition 4. 3.** *Soit $E$ un espace vectoriel. Une application $f : E \to E$ est dite affine si elle est la translatée d'une application linéaire, i.e. s'il existe $\mathbf{a} \in E$ et un endomorphisme $g \in \mathcal{L}(E)$ tels que*

$$f(\mathbf{x}) = \mathbf{a} + g(\mathbf{x}) \quad \forall \mathbf{x} \in E. \tag{4.6}$$

**Exemple.** Toute isométrie d'un espace vectoriel euclidien est affine (Th. 2.9).

**Application affine dans un espace affine.**

Soit $f : E \to E$ une application affine de l'espace vectoriel $E$ dans lui-même. Il existe $\mathbf{a} \in E$ et un endomorphisme $g \in \mathcal{L}(E)$ tels que (4.6) soit vérifiée. Comme $g$ est linéaire, on a $g(0) = 0$, donc $f(0) = \mathbf{a}$ et alors

$$f(\mathbf{x}) = f(0) + g(\mathbf{x}) \quad \forall \mathbf{x} \in E. \tag{4.7}$$

Si l'on note $X$ l'espace affine canoniquement associé à l'espace vectoriel $E$, un vecteur $\mathbf{x}$ est un point $M$ de $X$, le vecteur nul 0 est l'origine canonique $O$ et $\mathbf{OM} = \mathbf{x} \in E$. $f$ est une application de l'espace affine $X$ dans lui-même et l'équation (4.7) s'écrit

$$f(M) = f(O) + g(\mathbf{OM}) \quad \forall M \in X. \tag{4.8}$$

On constate donc que l'application $f : X \to X$ est affine au sens suivant.

**Définition 4. 4.** *Soit $X$ un espace affine associé à l'espace vectoriel $E$. Une application $f : X \to X$ est dite affine s'il existe un point $P \in X$ et un endomorphisme $g \in \mathcal{L}(E)$ tels que*

$$f(M) = f(P) + g(\mathbf{PM}) \quad \forall M \in X. \tag{4.9}$$

Si $f$ est une application affine de l'espace affine $X$ dans lui-même, *i.e.* s'il existe un point $P \in X$ et un endomorphisme $g \in \mathcal{L}(E)$ vérifiant l'équation (4.9) alors cette équation est encore vérifiée avec le *même* $g$ pour tout point $Q \in X$ à la place de $P$ :

$$f(M) = f(Q) + g(\mathbf{QM}) \quad \forall M, Q \in X. \tag{4.10}$$

En effet, pour tous $Q, M \in X$, on a

$$\begin{aligned} f(M) = f(P) + g(\mathbf{PM}) &= f(P) + g(\mathbf{PQ} + \mathbf{QM}) \\ &= f(P) + g(\mathbf{PQ}) + g(\mathbf{QM}) = f(Q) + g(\mathbf{QM}). \end{aligned}$$

De plus s'il existe un endomorphisme $g \in \mathcal{L}(E)$ vérifiant l'équation (4.9), il est unique puisque $g(\mathbf{PM}) = f(M) - f(P)$ $\forall M \in X$.

**Définition 4. 5.** *Soit $X$ un espace affine associé à l'espace vectoriel $E$ et $f : X \to X$ une application affine. L'unique endomorphisme $g \in \mathcal{L}(E)$ tel que*

$$f(M) = f(P) + g(\mathbf{PM}) \quad \forall P, M \in X. \tag{4.11}$$

*est appelé la partie linéaire de $f$.*

On notera que pour $P \in X$ fixé, l'équation (4.11) signifie que l'application $f$ est la composée des 3 applications

$$M \overset{\psi_P}{\longmapsto} \mathbf{PM} \overset{g}{\longmapsto} g(\mathbf{PM}) \overset{\varphi_{f(P)}}{\longmapsto} f(P) + g(\mathbf{PM})$$

où $\psi_P$ est la bijection (4.5) de $X$ sur $E$ définie par $\psi_P(M) = \mathbf{PM} \quad \forall M \in X$ et $\varphi_{f(P)}$ la bijection (4.4) de $E$ sur $X$ définie par $\varphi_{f(P)}(\mathbf{u}) = f(P) + \mathbf{u} \quad \forall \mathbf{u} \in E$. L'équation (4.11) s'écrit donc aussi :

$$f = \varphi_{f(P)} \circ g \circ \psi_P \quad \forall P \in X\,. \tag{4.12}$$

**Équivalence des deux définitions dans le cas de l'espace affine canoniquement associé à un espace vectoriel.**

Si $X$ est l'espace affine canoniquement associé à l'espace vectoriel $E$, et si $f$ est une application affine de $X$ dans lui-même, l'équation (4.10) peut être appliquée avec pour $Q$ l'origine canonique $O$ de $X$, et l'on obtient alors l'équation (4.8), avec $g$ la partie linéaire de $f$. Donc (4.6) est vérifiée avec $\mathbf{a} = f(0)$ et l'application $f$ de l'espace vectoriel $E$ dans lui même est affine.

Lorsque $X$ est l'espace affine canoniquement associé à l'espace vectoriel $E$, une application $f$ de $X$ dans lui-même est donc affine si et seulement si l'application $f$ de l'espace vectoriel $E$ dans lui-même est affine.

### 4.1.5 Groupe affine.

**Proposition 4. 2.** *Soit $X$ un espace affine associé à l'espace vectoriel $E$.*
*(i) Soit $f : X \to X$ une application affine et $g \in \mathcal{L}(E)$ sa partie linéaire. Alors $f$ est une bijection de $X$ sur $X$ si et seulement si $g$ est une bijection de $E$ sur $E$.*
*(ii) L'ensemble $GA(X)$ des bijections affines de $X$ sur $X$ est un groupe pour la loi $\circ$.*

**Démonstration.**
(i) Par définition de $g$ ,

$$f(M) = f(P) + g(\mathbf{PM}) \quad \forall P, M \in X.$$

On a vu (équation (4.12)) que cela s'écrit encore

$$f = \varphi_{f(P)} \circ g \circ \psi_P.$$

Si $g$ est bijective, alors $f$ est la composée de 3 bijections, donc est une bijection. Réciproquement,

$$g = (\varphi_{f(P)})^{-1} \circ f \circ (\psi_P)^{-1} = \psi_{f(P)} \circ f \circ \varphi_P.$$

Si $f$ bijective, $g$ est la composée de 3 bijections, donc est une bijection.
(ii) On a $GA(X) \subset \mathrm{Bij}\,(X)$. Montrons que c'est un sous-groupe de $\mathrm{Bij}\,(X)$.
• $Id_X \in GA(X)$ est clair. La partie linéaire de $Id_X$ est $Id_E$.
• $f_1 \circ f_2 \in GA(X)\ \forall\, f_1, f_2 \in GA(X)$. On a pour tous $P_1, P_2 \in X$ en notant $g_1, g_2$ les parties linéaires de $f_1, f_2$ respectivement :

$$\begin{aligned} f_1 &= \varphi_{f_1(P_1)} \circ g_1 \circ \psi_{P_1} \\ f_2 &= \varphi_{f_2(P_2)} \circ g_2 \circ \psi_{P_2}. \end{aligned}$$

Soit $P \in X$ quelconque et prenons $P_2 = P$ et $P_1 = f_2(P)$. Alors

$$\begin{aligned} f_1 \circ f_2 &= \varphi_{f_1(P_1)} \circ g_1 \circ \psi_{P_1} \circ \varphi_{f_2(P_2)} \circ g_2 \circ \psi_{P_2} \\ &= \varphi_{f_1(f_2(P))} \circ g_1 \circ \psi_{f_2(P)} \circ \varphi_{f_2(P)} \circ g_2 \circ \psi_P \\ &= \varphi_{f_1(f_2(P))} \circ g_1 \circ g_2 \circ \psi_P \\ &\phantom{=} (\text{car } \psi_{f_2(P)} \circ \varphi_{f_2(P)} = Id_E) \\ &= \varphi_{(f_1 \circ f_2)(P)} \circ (g_1 \circ g_2) \circ \psi_P. \end{aligned}$$

Cela s'écrit encore

$$(f_1 \circ f_2)(M) = (f_1 \circ f_2)(P) + (g_1 \circ g_2)(\mathbf{PM})\ \forall M \in X,$$

donc $f_1 \circ f_2$ est affine et sa partie linéaire est $g_1 \circ g_2$.
• $f^{-1} \in GA(X)\ \forall f \in GA(X)$. Soit $g$ la partie linéaire de $f$. On a pour $P \in X$

$$f = \varphi_{f(P)} \circ g \circ \psi_P$$

et on sait que $g$ est bijective car $f$ est bijective. Alors

$$f^{-1} = (\psi_P)^{-1} \circ g^{-1} \circ (\varphi_{f(P)})^{-1} = \varphi_P \circ g^{-1} \circ \psi_{f(P)}.$$

En posant $P = f^{-1}(Q)$, on a pour tout $Q \in X$

$$f^{-1} = \varphi_{f^{-1}(Q)} \circ g^{-1} \circ \psi_Q$$

donc $f^{-1}$ est affine et sa partie linéaire est $g^{-1}$. □

**Définition 4. 6.** *Soit $X$ un espace affine associé à l'espace vectoriel $E$. Le groupe $GA(X)$ est appelé groupe affine de $X$.*

### 4.1.6 Isométries d'un espace affine euclidien.

**Définition 4. 7.** *Soit $X$ un espace affine associé à l'espace vectoriel $E$ et $f : X \to X$ une application de $X$ dans lui-même. On dit que $f$ est une isométrie de $X$ si elle vérifie la condition*

$$\|\mathbf{f(A)\,f(B)}\| = \|\mathbf{AB}\| \quad \forall A, B \in X. \tag{4.13}$$

Dans le cas particulier où $X$ est l'espace affine canoniquement attaché à l'espace vectoriel $E$, la définition ci-dessus n'est autre que la définition 2.10.

On notera qu'une application *affine* $f : X \to X$ d'un espace affine $X$ associé à l'espace vectoriel $E$ est une isométrie de l'espace affine $X$ si et seulement si sa partie linéaire $g$ est un endomorphisme isométrique de $E$.

**Théorème 4. 1.** *Toute isométrie d'un espace affine euclidien $X$ est une bijection affine de $X$ sur $X$.*

**Démonstration.**

Soit $X$ un espace affine associé à l'espace vectoriel $E$ et $f : X \to X$ une isométrie de l'espace affine $X$. Montrons d'abord que $f$ est affine. Fixons une origine $O \in X$. Si $f$ était affine de partie linéaire $g$, on aurait $f(M) = f(O) + g(\mathbf{OM})\quad \forall M \in X$. Introduisons donc l'application $\hat{f} : E \to E$ définie par

$$f(M) = f(O) + \hat{f}(\mathbf{OM}) \quad \forall M \in X, \tag{4.14}$$

ce qui s'écrit encore

$$\hat{f}(\mathbf{OM}) = \mathbf{f(O)\,f(M)} \quad \forall M \in X. \tag{4.15}$$

On a pour tous $A, B \in X$ :

$$\mathbf{f(A)\,f(B)} = \mathbf{f(O)\,f(B)} - \mathbf{f(O)\,f(A)} = \hat{f}(\mathbf{OB}) - \hat{f}(\mathbf{OA}).$$

L'équation (4.13) s'écrit donc puisque $\mathbf{AB} = \mathbf{OB} - \mathbf{OA}$

$$\|\hat{f}(\mathbf{OB}) - \hat{f}(\mathbf{OA})\| = \|\mathbf{OB} - \mathbf{OA}\| \quad \forall A, B \in X$$

ou encore

$$\|\hat{f}(\mathbf{x}) - \hat{f}(\mathbf{y})\| = \|\mathbf{x} - \mathbf{y}\| \quad \forall \mathbf{x}, \mathbf{y} \in E.$$

Cela signifie que $\hat{f}$ est une isométrie de l'espace vectoriel euclidien $E$. Or on sait que toute isométrie de l'espace vectoriel euclidien $E$ est une application affine de $E$ (Th. 2.10). Donc il existe $\mathbf{a} \in E$ et $g \in \mathcal{L}(E)$ tels que $\hat{f} = \mathbf{a} + g$. Prenant $M = O$ dans (4.15) on a $\hat{f}(0) = 0$. Mais $g$ étant linéaire, on a aussi $g(0) = 0$. Donc $\mathbf{a} = 0$ et $\hat{f} = g$. Donc $\hat{f}$ est linéaire. Alors d'après (4.14), $f$ est affine et sa partie linéaire est $\hat{f}$.

Montrons maintenant que $f$ est une bijection de $X$ sur $X$. $f$ étant affine, on sait d'après la Prop.4.2 que $f$ est une bijection de $X$ sur $X$ si et seulement si sa partie linéaire est une bijection de $E$ sur $E$. Or la partie linéaire de $f$ est un endomorphisme isométrique de $E$ et donc c'est une bijection de $E$ sur $E$. D'où le résultat. □

### 4.1.7 Sous-espace affine.

**Proposition 4. 3.** *Soit $X$ un espace affine associé à l'espace vectoriel $E$, $Y$ un sous-ensemble de $X$ et $F$ un sous-espace vectoriel de $E$. Les conditions suivantes sont équivalentes.*

*(i) Il existe $A \in Y$ tel que $F = \{\mathbf{AM};\ M \in Y\}$.*

*(ii) Pour tout $B \in Y$ on a $F = \{\mathbf{BM};\ M \in Y\}$.*

*(iii) Il existe $A \in Y$ tel que $Y = A + F$.*

*(iv) Pour tout $B \in Y$ on a $Y = B + F$.*

*(v) $Y$ est une orbite de $X$ pour l'action sur $X$ du groupe additif de $F$ obtenue par restriction de l'action du groupe additif de $E$.*

**Démonstration.**

(ii) $\Rightarrow$ (i) est trivial.

(i) $\Rightarrow$ (ii). Soit $B \in Y$. D'après (i), $\mathbf{AB} \in F$ et $\mathbf{AM} \in F$ pour tout $M \in Y$. Comme $F$ est un sous-espace vectoriel, on a donc $\mathbf{BM} = \mathbf{AM} - \mathbf{AB} \in F$ pour tout $M \in Y$, ce qui montre que $\{\mathbf{BM};\ M \in Y\} \subset F$. Soit maintenant $\mathbf{t} \in F$. D'après (i), il existe $M \in Y$ tel que $\mathbf{t} = \mathbf{AM}$. Soit $N = M + \mathbf{AB}$. On a $\mathbf{AN} = \mathbf{AM} + \mathbf{MN} = \mathbf{t} + \mathbf{AB} \in F$, donc $N \in Y$ d'après (i). Or $\mathbf{BN} = \mathbf{AN} - \mathbf{AB} = \mathbf{t}$. Donc $\mathbf{t} \in \{\mathbf{BM};\ M \in Y\}$. Comme $\mathbf{t} \in F$ est arbitraire, $F \subset \{\mathbf{BM};\ M \in Y\}$, d'où l'égalité.

(i) $\Leftrightarrow$ (iii) et (ii) $\Leftrightarrow$ (iv). Pour tout point $A \in X$ on a

$$A + F = \{M \in E;\ \exists \mathbf{t} \in F,\ M = A + \mathbf{t}\} = \{M \in E;\ \mathbf{AM} \in F\}.$$

Donc

$$\begin{aligned} Y = A + F &\Leftrightarrow Y = \{M \in E;\ \mathbf{AM} \in F\} \\ &\Leftrightarrow \forall M \in E,\ (M \in Y \Leftrightarrow \mathbf{AM} \in F) \\ &\Leftrightarrow F = \{\mathbf{AM};\ M \in Y\}. \end{aligned}$$

Cela prouve simultanément (i) $\Leftrightarrow$ (iii) et (ii) $\Leftrightarrow$ (iv)
(iii) $\Leftrightarrow$ (v). Résulte de la définition de l'orbite par $F$ d'un point $A$. □

**Définition 4. 8.** *Soit $X$ un espace affine associé à l'espace vectoriel $E$. Un sous-ensemble $Y$ de $X$ est appelé sous-espace affine de $X$ s'il existe un sous-espace vectoriel $F$ vérifiant les conditions équivalentes de la Proposition 4.3. Le sous-espace vectoriel $F$ (qui est unique d'après la Proposition 4.3) est appelé direction du sous-espace affine $Y$ de $X$.*

## 4.2 Barycentre.

### 4.2.1 Définition du barycentre.

**Proposition 4. 4.** *Soit $X$ un espace affine, $O$ un point de $X$ pris comme origine, $A_1, \dots, A_n$ des points de $X$ et $\lambda_1, \dots, \lambda_n \in K$ des scalaires tels que $\sum_{i=1}^n \lambda_i \neq 0$.*
*(i) Soit $G$ le point de $X$ tel que*

$$\mathbf{OG} = \frac{\sum_{i=1}^n \lambda_i \mathbf{OA_i}}{\sum_{i=1}^n \lambda_i}. \tag{4.16}$$

*$G$ ne dépend pas du point $O$ utilisé comme origine.*
*(ii) $G$ est le seul point de $X$ tel que*

$$\sum_{i=1}^n \lambda_i \mathbf{GA_i} = 0. \tag{4.17}$$

**Démonstration.**

(i) Notons $S = \sum_{i=1}^n \lambda_i$. Si $O'$ est un autre point pris comme origine et si $G'$ est le point défini par $\mathbf{O'G'} = \frac{1}{S}\sum_{i=1}^n \lambda_i \mathbf{O'A_i}$, on aura

$$\mathbf{O'G'} = \frac{1}{S}\sum_{i=1}^n \lambda_i(\mathbf{O'O} + \mathbf{OA_i}) = \mathbf{O'O} + \frac{1}{S}\sum_{i=1}^n \lambda_i \mathbf{OA_i} = \mathbf{O'O} + \mathbf{OG} = \mathbf{O'G}$$

donc $G = G'$.
(ii) Il suffit de noter que pour un point $M \in X$ quelconque, on a

$$\begin{aligned} \sum_{i=1}^n \lambda_i \mathbf{MA_i} &= \sum_{i=1}^n \lambda_i(\mathbf{OA_i} - \mathbf{OM}) \\ &= \sum_{i=1}^n \lambda_i \mathbf{OA_i} - S\,\mathbf{OM} \\ &= S\,\mathbf{OG} - S\,\mathbf{OM} \\ &= S(\mathbf{OG} - \mathbf{OM}) \\ &= S\,\mathbf{MG}. \end{aligned}$$

□

**Définition 4. 9.** *Le point $G$ défini par (4.16) est appelé le barycentre des points $A_1, \dots, A_n$ affectés des poids $\lambda_1, \dots, \lambda_n$ tels que $\sum_{i=1}^n \lambda_i \neq 0$.*

**Remarque.** On ne change pas le barycentre $G$ en multipliant tous les poids par un même facteur constant $\alpha \neq 0$. En particulier, divisant chaque $\lambda_i$ par la somme des poids $S = \sum_{i=1}^{n} \lambda_i$, on est ramené au cas où $\sum_{i=1}^{n} \lambda_i = 1$. Dans la suite, on supposera le plus souvent que $\sum_{i=1}^{n} \lambda_i = 1$. Lorsque tous les poids sont égaux, on dit que $G$ est *l'isobarycentre* de points $A_1, \ldots, A_n$.

**Exemple.** Soient $A, B$ deux points de $X$. L'ensemble des barycentres des points $A$ et $B$ affectés des poids $\lambda$ et $1-\lambda$ respectivement, avec $0 \leqslant \lambda \leqslant 1$, est appelé le segment $[A, B]$. Si $O$ est un point fixé de $X$, on a donc

$$[A, B] = \{M \in X;\ \exists \lambda, 0 \leqslant \lambda \leqslant 1,\ \mathbf{OM} = \lambda \mathbf{OA} + (1-\lambda)\mathbf{OB}\}.$$

**Proposition 4. 5.** *Soit $X$ un espace affine associé à l'espace vectoriel $E$, $A_1, \ldots, A_n$ des points de $X$ et $F = vect\,(\mathbf{A_1A_2}, \ldots, \mathbf{A_1A_n})$ le sous-espace vectoriel de $E$ engendré par les vecteurs $\mathbf{A_1A_2}, \ldots, \mathbf{A_1A_n}$. Le sous-espace affine $Y = A_1 + F$ est l'ensemble de tous les barycentres formés avec les points $A_1, \ldots, A_n$.*

**Démonstration.**
Par définition de $Y$, un point $M$ de $X$ appartient à $Y$ si et seulement si il existe des scalaires $\lambda_2, \ldots, \lambda_n$ tel que $\mathbf{A_1M} = \sum_{i=2}^{n} \lambda_i\, \mathbf{A_1A_i}$. Mais en posant $\lambda_1 = 1 - \sum_{i=2}^{n} \lambda_i$, cela équivaut à dire que $\mathbf{A_1M} = \sum_{i=1}^{n} \lambda_i\, \mathbf{A_1A_i}$, *i.e.* $M$ est le barycentre des points $A_1, \ldots, A_n$ affectés des poids $\lambda_1, \ldots, \lambda_n$. □

### 4.2.2 Associativité des barycentres.

**Théorème 4. 2.** *Soit $G$ le barycentre des points $A_1, \ldots, A_n$ affectés des poids $\lambda_1, \ldots, \lambda_n$ tels que $\sum_{i=1}^{n} \lambda_i \neq 0$. Soit $q$ tel que $1 < q < p$ et $\sum_{i=1}^{q} \lambda_i \neq 0$. Si $G'$ désigne le barycentre des points $A_1, \ldots, A_q$ affectés des poids $\lambda_1, \ldots, \lambda_q$, alors $G$ est le barycentre des points $G', A_{q+1}, \ldots, A_n$ affectés des poids $\lambda, \lambda_{q+1}, \ldots, \lambda_n$, avec $\lambda = \sum_{i=1}^{q} \lambda_i$.*

**Démonstration.**
Soit $O$ est un point de $X$ fixé, $S = \sum_{i=1}^{n} \lambda_i$ et $\lambda = \sum_{i=1}^{q} \lambda_i$. On a :

$$\begin{aligned} S\,\mathbf{OG} &= \sum_{i=1}^{n} \lambda_i \mathbf{OA_i} \\ &= \sum_{i=1}^{q} \lambda_i \mathbf{OA_i} + \sum_{i=q+1}^{n} \lambda_i \mathbf{OA_i} \\ &= \lambda\,\mathbf{OG'} + \sum_{i=q+1}^{n} \lambda_i \mathbf{OA_i} \end{aligned}$$

d'où le résultat puisque $S = \lambda + \sum_{i=q+1}^{n} \lambda_i$. □

### 4.2.3 Image du barycentre par une application affine.

**Théorème 4. 3.** *Soit $X$ un espace affine et $f : X \to X$ une application affine. Soit $G$ le barycentre des points $A_1, \ldots, A_n$ affectés des poids $\lambda_1, \ldots, \lambda_n$ tels que $S = \sum_{i=1}^{n} \lambda_i \neq 0$. Alors $f(G)$ est le barycentre des points $f(A_1), \ldots, f(A_n)$ affectés des poids $\lambda_1, \ldots, \lambda_n$.*

**Démonstration.**

On peut supposer $S = 1$. Si $O$ est un point de $X$ fixé, on a $\mathbf{OG} = \sum_{i=1}^n \lambda_i \mathbf{OA_i}$. Comme $f$ est affine, il existe $g \in \mathcal{L}(E)$ telle que

$$f(M) = f(O) + g(\mathbf{OM}) \quad \forall M \in X.$$

Alors

$$f(G) = f(O) + g\left(\sum_{i=1}^n \lambda_i \mathbf{OA_i}\right) = f(O) + \sum_{i=1}^n \lambda_i g(\mathbf{OA_i})$$

donc

$$f(G) - f(O) = \sum_{i=1}^n \lambda_i g(\mathbf{OA_i}) = \sum_{i=1}^n \lambda_i (f(A_i) - f(O))$$

ou encore

$$\mathbf{f(O)f(G)} = \sum_{i=1}^n \lambda_i \mathbf{f(O)f(A_i)}.$$

Cela signifie que $f(G)$ est le barycentre des points $f(A_1), \ldots, f(A_n)$ affectés des poids $\lambda_1, \ldots, \lambda_n$. □

**Exemple.** L'image par une application affine $f$ d'un segment $[A, B]$ est le segment $[f(A), f(B)]$.

### 4.2.4 Applications.

**Isobarycentre des sommets d'un triangle.**

Soient $A, B, C$ les 3 sommets d'un triangle. Soit $H$ le barycentre des points $B, C$ affectés des poids $1, 1$. L'isobarycentre $G$ des points $A, B, C$ affectés des poids $1, 1, 1$ est aussi le barycentre de $H$ et $A$ affectés respectivement des poids 2 et 1. Le point $H$ est le milieu du segment $[B, C]$. La droite $AH$ est la médiane du triangle issue de $A$. Le point $G$ vérifie $\mathbf{GA} + 2\mathbf{GH} = 0$, *i.e.* $\mathbf{GA} = -2\mathbf{GH}$, donc $\mathbf{AG} = \frac{2}{3}\mathbf{AH}$. Permutant circulairement le rôle des points $A, B, C$, on voit que $G$ est le point de concours des 3 médianes du triangle, et est situé "au $\frac{2}{3}$ de leur longueur". Dans le cas particulier où le triangle est équilatéral, chaque médiane est aussi la médiatrice du segment correspondant et donc la hauteur issue du sommet opposé.

**Isobarycentre des sommets d'un tétraèdre régulier.**

Soient $A, B, C, D$ les 4 sommets d'un tétraèdre régulier, *i.e.* dont les faces sont des triangles équilatéraux (voir Fig. 4.1). Soient $H, K$ les milieux des segments $[C, D], [B, C]$ respectivement. L'isobarycentre $G'$ des points $B, C, D$ est le point d'intersection des droites $BH$ et $DK$. Nous allons d'abord constater que la droite $AG'$ est perpendiculaire au plan $BCD$.

L'ensemble des points $M \in \mathbb{R}^3$ qui sont équidistants de deux points fixés $P, Q$ est un plan perpendiculaire au segment $[PQ]$ en son milieu $I$, appelé *plan médiateur de* $[P, Q]$. On a en effet

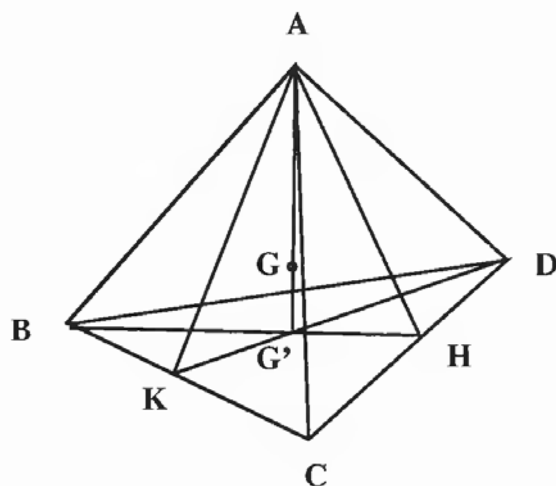



FIG. 4.1: *Isobarycentre des sommets du tétraèdre régulier.*

$$\begin{aligned}\|\mathbf{MP}\|^2 = \|\mathbf{MQ}\|^2 &\Leftrightarrow \|\mathbf{MI}+\mathbf{IP}\|^2 = \|\mathbf{MI}+\mathbf{IQ}\|^2\\ &\Leftrightarrow (\mathbf{MI}|\mathbf{IP}) = (\mathbf{MI}|\mathbf{IQ})\\ &\Leftrightarrow (\mathbf{MI}|\mathbf{IP}) = (\mathbf{MI}|-\mathbf{IP})\\ &\Leftrightarrow (\mathbf{MI}|\mathbf{IP}) = 0\end{aligned}$$

d'où le résultat puisque **IP** est un vecteur directeur de la droite $PQ$.

Le plan médiateur de $[C, D]$ est le plan $ABH$. Tout vecteur de ce plan est donc perpendiculaire à la droite $CD$. De même, le plan médiateur de $[B, C]$ est le plan $ADK$ et tout vecteur de ce plan est perpendiculaire à la droite $BC$. L'intersection de plans $ABH$ et $ADK$ contient les points $A$ et $G'$, donc c'est la droite $AG'$. Ainsi le vecteur $\mathbf{AG'}$ est orthogonal au vecteur **BC** et au vecteur **CD**. Il est donc orthogonal au plan $BCD$.

Maintenant, l'isobarycentre $G$ des 4 points $A, B, C, D$ est le barycentre du point $G'$ affecté du poids 3 et du point $A$ affecté du poids 1. Le point $G$ vérifie $\mathbf{GA} + 3\mathbf{GG'} = 0$, *i.e.* $\mathbf{GA} = -3\mathbf{GG'}$. Permutant circulairement le rôle des points $A, B, C, D$ on voit que $G$ est le point de concours des 4 hauteurs du tétraèdre, et est situé "au $\frac{3}{4}$ de leur longueur".

**Isobarycentre des sommets d'un carré.**

Soient $A, B, C, D$ les 4 sommets d'un carré. L'isobarycentre des points $A, B, C, D$ est le centre $O$ du carré puisque $\mathbf{OC} = -\mathbf{OA}, \mathbf{OD} = -\mathbf{OB}$ implique $\mathbf{OA} + \mathbf{OB} + \mathbf{OC} + \mathbf{OD} = 0$.

### 4.2.5 Parties convexes.

**Définition 4. 10.** *Soit $X$ un espace affine. Une partie $Y$ de $X$ est dite convexe si pour tous $A, B \in Y$ le segment $[A, B]$ est inclus dans $Y$.*

**Proposition 4. 6.** *Une partie $Y$ d'un espace affine $X$ est convexe si et seulement si tout barycentre de points de $Y$ avec des poids $\geqslant 0$ appartient à $Y$.*

**Démonstration.**

*Condition nécessaire.* Supposons $Y$ convexe, et soit $G$ le barycentre d'une famille de points $A_1, \ldots, A_n$ de $Y$ affectés des poids $\lambda_1, \ldots, \lambda_n$ avec $\lambda_i \geqslant 0 \ \forall i$ et $\sum_{i=1}^n \lambda_i \neq 0$. On peut supposer $\sum_{i=1}^n \lambda_i = 1$ et $\lambda_i > 0 \ \forall i$. Si $O$ est un point de $X$ fixé, on a donc $\mathbf{OG} = \sum_{i=1}^n \lambda_i \mathbf{OA_i}$. Si $n = 2$, $G \in Y$ par définition de la convexité.

On raisonne maintenant par récurrence sur $n$ en supposant le résultat vrai pour le niveau $n = p$. Pour $n = p+1$, on aura

$$\mathbf{OG} = \sum_{i=1}^{p+1} \lambda_i \mathbf{OA_i} = \sum_{i=1}^{p} \lambda_i \mathbf{OA_i} + \lambda_{p+1}\mathbf{OA_{p+1}} = \lambda \mathbf{OG_p} + (1-\lambda)\mathbf{OA_{p+1}}$$

où $G_p$ désigne le barycentre des points $A_1, \ldots, A_p$ affectés des poids $\lambda_1, \ldots, \lambda_p$ avec $\lambda = \sum_{i=1}^p \lambda_i$. D'après l'hypothèse de récurrence, $G_p \in Y$. Comme $0 \leqslant \lambda \leqslant 1$, $G \in [G_p, A_{p+1}]$, donc $G \in Y$ par convexité de $Y$.
*Condition suffisante.* Immédiate puisque la convexité correspond au cas particulier de deux points. □

**Corollaire.** *Pour toute partie $Y$ de $X$, l'ensemble de tous les barycentres de points de $Y$ avec des poids $\geqslant 0$ est le plus petit ensemble convexe contenant $Y$. On l'appelle enveloppe convexe de $Y$.*

**Démonstration.**

Soit $\mathcal{C}$ l'ensemble de tous les barycentres de points de $Y$ avec des poids $\geqslant 0$, *i.e.* l'ensemble des points $M \in X$ pour lesquel il existe $n \in \mathbb{N}$, $n \geqslant 2$, des points $A_1, \cdots, A_n \in Y$, et des poids $\lambda_1, \cdots, \lambda_n \geqslant 0$ tels que $M$ soit le barycentre des points $A_1, \cdots, A_n$ affectés des poids $\lambda_1, \cdots, \lambda_n \geqslant 0$. D'après la Prop. 4.6, toute partie convexe de $X$ contenant $Y$ contient aussi $\mathcal{C}$. Or $\mathcal{C}$ est convexe. En effet, si $M, M' \in \mathcal{C}$, $O \in X$ étant fixé, on peut supposer, quitte à rajouter éventuellement des poids nuls, que l'on a avec le même $n$ $\mathbf{OM} = \sum_{i=1}^n \lambda_i \mathbf{OA_i}$ et $\mathbf{OM'} = \sum_{i=1}^n \lambda'_i \mathbf{OA_i}$ où $A_i \in Y$, $\lambda_i, \lambda'_i \geqslant 0 \ \forall i$, $\sum_{i=1}^n \lambda_i > 0, \sum_{i=1}^n \lambda'_i > 0$. Alors pour $0 \leqslant \lambda \leqslant 1$, $\lambda\mathbf{OM} + (1-\lambda)\mathbf{OM'} = \sum_{i=1}^n (\lambda\lambda_i + (1-\lambda)\lambda'_i)\mathbf{OA_i}$ avec

$$\sum_{i=1}^n (\lambda\lambda_i + (1-\lambda)\lambda'_i) = \lambda(\sum_{i=1}^n \lambda_i) + (1-\lambda)\sum_{i=1}^n \lambda'_i > 0,$$

donc $[M, M'] \subset \mathcal{C}$. Ainsi $\mathcal{C}$ est convexe, et donc c'est la plus petite partie convexe contenant $Y$. □

### 4.2.6 Points extrêmaux d'une partie convexe.

**Définition 4. 11.** *Soit $X$ un espace affine et $Y$ une partie convexe de $X$. Un point $A \in Y$ est dit extrêmal s'il n'existe aucun segment ouvert de $X$ contenant $A$, i.e. si $A$ n'est pas barycentre de 2 points distincts $P, Q \in Y$, $P \neq Q$ avec des poids $\lambda, 1-\lambda$ et $0 < \lambda < 1$.*

**Exemples.**

(i) Soit $\Delta$ une droite affine du plan affine euclidien $E$. Si l'on introduit un repère affine $(O, (\mathbf{e_1}, \mathbf{e_2}))$, il existe $u, v, w \in \mathbb{R}$ tels que $\Delta$ a pour équation $\varphi(M) = 0$, en notant $\varphi(M) = ux + vy + w$ où $(x, y)$ sont les coordonnées du point $M$. $\Delta$ définit deux *demi-plans* $\{M \,;\, \varphi(M) \geqslant 0\}$ et $\{M \,;\, \varphi(M) \leqslant 0\}$. Il est immédiat que chacun de ces demi-plans est convexe.

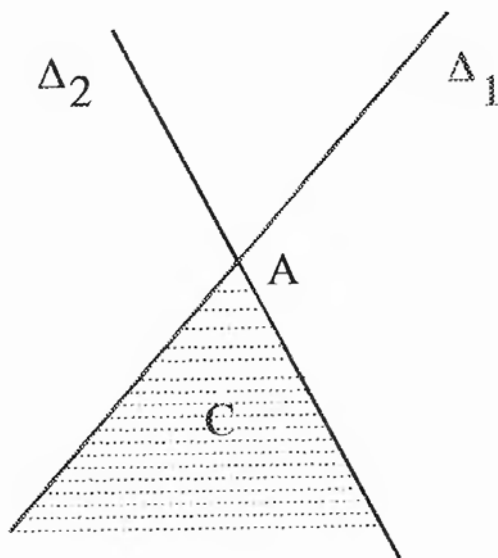



FIG. 4.2: *Point extrêmal de C.*

Soient maintenant deux droites affines $\Delta_1$ et $\Delta_2$ concourantes distinctes, et $A$ leur point d'intersection. Les équations de $\Delta_1$ et $\Delta_2$ sont respectivement $\varphi_1(M) = 0$ et $\varphi_2(M) = 0$, avec $\varphi_1(M) = u_1x + v_1y + w_1$ et $\varphi_2(M) = u_2x + v_2y + w_2$ $(u_1, v_1, w_1, u_2, v_2, w_2 \in \mathbb{R})$. Considérons l'intersection $C$ de deux demi-plans $\Pi_1$ et $\Pi_2$ définis respectivement par $\Delta_1$ et $\Delta_2$. On peut supposer (quitte à changer éventuellement $\varphi_i$ en $-\varphi_i$ si besoin pour $i = 1$ et/ou $i = 2$) $\Pi_i = \{M \in E\,;\, \varphi_i(M) \geqslant 0\}$ $\forall\, i = 1, 2$. Comme $C$ est l'intersection de deux convexes, $C$ est convexe. Nous allons voir que $A$ est l'unique point extrêmal de $C$ (voir Fig. 4.2).

Vérifions que $A$ est un point extrêmal de $C$. On a déjà $A \in C$. Si $A$ n'était pas un point extrêmal de $C$, il existerait deux points distincts $P, Q \in C$ et $\lambda \in ]0,1[$ tels que $\mathbf{OA} = \lambda\mathbf{OP} + (1-\lambda)\mathbf{OQ}$ ou encore en notant $(x_M, y_M)$ les coordonnées d'un point $M$ : $x_A = \lambda x_P + (1-\lambda)x_Q,\ y_A = \lambda y_P + (1-\lambda)y_Q$. Alors pour $i = 1, 2$ on a $u_ix_A + v_iy_A + w_i = \lambda(u_ix_P + v_iy_P + w_i) + (1-\lambda)(u_ix_Q + v_iy_Q + w_i)$, *i.e.*

$$\varphi_i(A) = \lambda\varphi_i(P) + (1-\lambda)\varphi_i(Q). \tag{4.18}$$

Puisque $A \in \Delta_i$ on a $\varphi_i(A) = 0$. Mais $P, Q \in \Pi_i$ donc $\varphi_i(P), \varphi_i(Q) \geqslant 0$. Comme $0 < \lambda < 1$, la condition (4.18) implique donc $\varphi_i(P) = \varphi_i(Q) = 0$. On obtient ainsi $P, Q \in \Delta_1 \cap \Delta_2 = \{A\}$ ce qui est contradictoire car $P$ et $Q$ sont distincts. Donc $A$ est un point extrêmal de $C$.

Montrons maintenant qu'un point $B \in C$ distinct de $A$ n'est pas un point extrêmal de $C$. 3 cas sont possibles :

• $\varphi_1(B) > 0, \varphi_2(B) > 0$ : Comme $\varphi_1$ et $\varphi_2$ sont des fonctions continues du couple $(x, y)$ des coordonnées de $M$, il existe $r_1, r_2 > 0$ tels que $\|\mathbf{BM}\| < r_1$ implique $\varphi_1(M) > 0$ et $\|\mathbf{BM}\| < r_2$ implique $\varphi_2(M) > 0$. Soit $r = \inf(r_1, r_2)$. Alors quel que soit le vecteur normé $\mathbf{u}$, les deux points $P = B + \frac{r}{2}\mathbf{u}$ et $Q = B - \frac{r}{2}\mathbf{u}$ appartiennent à $C$. Or $B$ est le milieu du segment $[P, Q]$ donc $B$ n'est pas un point extrêmal de $C$.

• $\varphi_1(B) > 0, \varphi_2(B) = 0$ : il existe $r_1$ tel que $\|\mathbf{BM}\| < r_1$ implique $\varphi_1(M) > 0$. Alors si $\mathbf{u_2}$ désigne un vecteur directeur normé de $\Delta_2$, les deux points $P = B + \frac{r_1}{2}\mathbf{u_2}$ et $Q = B - \frac{r_1}{2}\mathbf{u_2}$ appartiennent à $C$. Or $B$ est le milieu du segment $[P, Q]$ donc $B$ n'est pas un point extrêmal de $C$.

• $\varphi_1(B) = 0, \varphi_2(B) > 0$ : analogue au cas précédent.

(ii) Soit $P$ un polygone convexe à $n$ côtés, et $A_0, A_1, \cdots, A_{n-1}$ ses sommets. $P$ est l'intersection de $n$ demi-plans $\Pi_0, \Pi_1, \cdots, \Pi_{n-1}$ définis respectivement par les droites affines $\Delta_0, \Delta_1, \cdots, \Delta_{n-1}$ passant par les couples de sommets consécutifs $\{A_0, A_1\}, \cdots, \{A_{n-1}, A_0\}$ respectivement :

$$P = \bigcap_{i=0}^{n-1} \Pi_i\,.$$

Considérons le point $A_0$ de $P$. C'est d'après l'exemple (i) un point extrêmal de $\Pi_0 \cap \Pi_1$. Donc c'est *a fortiori* un point extrêmal de $P$. De même $A_1, \cdots, A_{n-1}$ sont des points extrêmaux de $P$. Tous les sommets de $P$ sont donc des points extrêmaux de $P$. Comme dans l'exemple (ii), on pourrait démontrer, mais nous ne le ferons pas, qu'un point de $P$ qui n'est pas un sommet n'est pas extrêmal. On conclut donc que l'ensemble des points extrêmaux de $P$ est l'ensemble $A_0, A_1, \cdots, A_{n-1}$ de ses sommets.

(iii) Soit $\Sigma$ un plan affine de l'espace affine euclidien à 3 dimensions. Si l'on introduit un repère affine $(O, (\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}))$, il existe $u, v, w, t \in \mathbb{R}$ tels que $\Sigma$ a pour équation $\varphi(M) = 0$, en notant $\varphi(M) = ux + vy + wz + t$ où $(x, y, z)$ sont les coordonnées du point $M$. $\Sigma$ définit deux *demi-espaces* $\{M\,;\, \varphi(M) \geqslant 0\}$ et $\{M\,;\, \varphi(M) \leqslant 0\}$. Il est immédiat que chacun de ces demi-espaces est convexe.

Soient $\Sigma_1, \Sigma_2, \Sigma_3$ 3 plans affines 2 à 2 non parallèles de l'espace affine euclidien à 3 dimensions. L' équation de $\Sigma_i$ $(1 \leqslant i \leqslant 3)$ est $\varphi_i(M) = 0$, avec $\varphi_i(M) = u_i x + v_i y + w_i z + t_i$, $(u_i, v_i, w_i, t_i \in \mathbb{R})$. Soient $\Delta_1 = \Sigma_1 \cap \Sigma_2$, $\Delta_2 = \Sigma_2 \cap \Sigma_3, \Delta_3 = \Sigma_3 \cap \Sigma_1$. On suppose $\Delta_1, \Delta_2, \Delta_3$ concourantes en un point $A$. Considérons l'intersection $C$ de 3 demi-espaces $\Phi_1, \Phi_2, \Phi_3$ définis respectivement par $\Sigma_1$, $\Sigma_2$ et $\Sigma_3$. On peut supposer (quitter à changer éventuellement $\varphi_i$ en $-\varphi_i$ si besoin pour $i = 1$ et/ou $i = 2$ et/ou $i = 3$) $\Phi_i = \{M \in E\,;\, \varphi_i(M) \geqslant 0\}$ $\forall i$ $1 \leqslant i \leqslant 3$. Comme $C$ est une intersection de convexes, $C$ est convexe. En reprenant exactement la même raisonnement que dans l'exemple (i) mais avec 3 composantes pour les vecteurs, on verrait que $A$ est un point extrêmal de $C$. On verrait également de même que $A$ est l'unique point extrêmal de $C$.

(iv) On appelle *polyèdre convexe* de l'espace affine euclidien à 3 dimensions toute partie $\Lambda$ non vide qui est l'intersection d'une suite *finie* de demi-espaces $\Phi_1, \cdots, \Phi_p$ définis respectivement par des plans affines $\Sigma_1, \cdots, \Sigma_p$ $(p \geqslant 4)$ :

$$\Lambda = \bigcap_{i=1}^{p} \Pi_i\,.$$

Les *faces* de $\Lambda$ sont les intersections de $\Lambda$ avec les divers plans. Les intersections de 2 (*resp :* au moins 3) faces sont les *arêtes* (*resp : sommets*) de $\Lambda$. Les faces sont des polygones convexes. D'après l'exemple (iii), les sommets de $\Lambda$ sont des points extrêmaux de $\Lambda$. Comme à l'exemple (ii), on pourrait démontrer, mais nous ne le ferons pas, que ce sont les seuls points extrêmaux de $\Lambda$.

**Proposition 4. 7.** *Soit $X$ un espace affine et $f : X \to X$ une application affine bijective. Soit $C$ une partie convexe de $X$ et $Extr(C)$ l'ensemble de ses points extrêmaux. Si $f(C) = C$, alors $f(Extr(C)) = Extr(C)$.*

**Démonstration.**

On peut supposer $Extr(C) \neq \emptyset$. Soit $P \in Extr(C)$, et $P' = f(P)$. Si $P' \notin Extr(C)$, $P'$ serait barycentre de deux points distincts $A', B' \in C$ avec des coefficients $\lambda$ et $1-\lambda$ ($\lambda \in ]0,1[$). Comme $f(C) = C$, et que $f$ est une bijection, on a aussi $f^{-1}(C) = C$. Or on sait que l'application réciproque d'une application affine bijective est affine. Donc $f^{-1}$ est affine. Par conséquent $P = f^{-1}(P')$ serait le barycentre des deux points distincts $A = f^{-1}(A'), B = f^{-1}(B') \in C$ avec les mêmes poids $\lambda$ et $1-\lambda$. Cela n'est pas possible puisque $P$ est un point extrêmal de $C$. Donc $P' \in Extr(C)$. $P \in Extr(C)$ étant arbitraire, $f(Extr(C)) \subset Extr(C)$. Maintenant cette inclusion est vraie pour toute bijection affine $f$ de $E$ laissant stable $C$. Donc aussi en particulier pour $f^{-1}$. Cela donne $f^{-1}(Extr(C)) \subset Extr(C)$ d'où, en composant par $f$, $Extr(C) \subset f(Extr(C))$. □

Signalons enfin le Théorème suivant (voir [17] Th. 10.4 ou [3] Chap. 2, par. 7, Th. 1 p. 107).

**Théorème 4. 4 (Krein-Milman).** *Soit $X$ un espace affine euclidien et $Y$ une partie convexe compacte non vide de $E$. Alors*
*(i) L'ensemble des points extrêmaux de $Y$ est non vide : $Extr(Y) \neq \emptyset$ ;*
*(ii) $Y$ est la fermeture de l'enveloppe convexe de l'ensemble de ses points extrêmaux.*

L'hypothèse sur $X$ faite ici est une simplification, car le Th. de Krein-Milman est valable pour des espaces plus généraux.

## 4.3 Centre de gravité des solides.

Dans cette section, on prend pour $X$ un espace affine euclidien de dimension 3, muni d'un repère affine orthonormé $(O, (\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}))$. Si un point $M \in X$ a pour coordonnées $(x, y, z)$, $\mathbf{OM} = x\,\mathbf{e_1} + y\,\mathbf{e_2} + z\,\mathbf{e_3}$. On identifie l'ensemble $X$ à l'espace vectoriel euclidien $\mathbb{R}^3$ en identifiant $M \in X$ ou le vecteur $\mathbf{OM}$ au vecteur $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$.

### 4.3.1 Longueur d'un arc.

**Arc géométrique de classe $C^1$.**

Un *arc paramétré de classe $C^1$* est un couple $(I, \gamma)$, où $I$ est un intervalle de $\mathbb{R}$ et $\gamma : t \mapsto \gamma(t)$ une application de classe $C^1$ de $I$ dans $\mathbb{R}^3$. Deux arcs paramétrés de classe $C^1$ $(I, \gamma)$ et $(J, \lambda)$ sont dits *équivalents* s'il existe un $C^1$-difféomorphisme $\varphi : I \to J$, *i.e.* une application bijective, dérivable à dérivée continue et telle que $\varphi'(t) \neq 0 \ \forall t \in I$, vérifiant

$$\gamma = \lambda \circ \varphi \tag{4.19}$$

On définit ainsi une relation d'équivalence sur l'ensemble des arcs paramétrés de classe $C^1$. Une classe d'équivalence est appelée *arc géométrique de classe $C^1$*. Etant donné un arc géométrique représenté par l'arc paramétré $(I, \gamma)$, tout autre arc paramétré équivalent est encore appelé *paramétrage admissible* de l'arc. Tous les paramétrages admissibles ont la même image, appelée *support* de l'arc géométrique. En imposant dans (4.19) la condition $\varphi'(t) > 0$ au lieu de $\varphi'(t) \neq 0$ on obtient la notion d'arc géométrique *orienté* et de paramétrage admissible de l'arc orienté.

**Longueur d'un arc.**

Soit $([a,b],\gamma)$ un arc paramétré de classe $C^1$ avec $a,b \in \mathbb{R}$, $a<b$. La longueur $L$ de l'arc est définie par

$$L = \int_a^b \|\gamma'(t)\|\,dt = \int_a^b \sqrt{x'^2 + y'^2 + z'^2}\,dt \tag{4.20}$$

si $\gamma(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}$. $L$ ne dépend que de l'arc géométrique et non du paramétrage admissible utilisé. En effet, soit $(J,\lambda)$ un autre paramétrage admissible vérifiant (4.19). $J = \varphi([a,b])$ est un intervalle fermé de la forme $[c,d]$, $c,d \in \mathbb{R}, c<d$. Comme $\varphi'(t) \neq 0\, \forall t \in [a,b]$, $\varphi'$ garde un signe constant sur $[a,b]$. Soit $\varepsilon = \pm 1$ le signe de $\varphi'$. Alors

$$\begin{aligned} L = \int_a^b \|\gamma'(t)\|\,dt &= \int_a^b \|\lambda'(\varphi(t))\varphi'(t)\|\,dt = \int_a^b \|\lambda'(\varphi(t))\|\,|\varphi'(t)|\,dt \\ &= \varepsilon \int_a^b \|\lambda'(\varphi(t))\|\,\varphi'(t)\,dt = \int_c^d \|\lambda'(u)\|\,du. \end{aligned}$$

**Abscisse curviligne.**

L'application

$$t \mapsto s(t) = \int_a^t \|\gamma'(v)\|\,dv$$

est dérivable sur $[a,b]$, de dérivée $s'(t) = \|\gamma'(t)\|$, donc c'est une application croissante de $[a,b]$ sur $[s(a),s(b)]$. Si l'arc est sans points stationnaires, *i.e.* $\gamma'(t) \neq 0\, \forall t \in [a,b]$, on a $s'(t) > 0\ \forall t \in [a,b]$. Dans ce cas, l'application $t \mapsto s(t)$ de $[a,b]$ sur $[s(a),s(b)]$ est un $C^1$-difféomorphisme croissant. Si l'on désigne alors par $\varphi : s \mapsto t(s)$ l'application réciproque, et si l'on pose $\lambda = \gamma \circ \varphi$, $([s(a),s(b)],\lambda)$ est un paramétrage admissible de l'arc géométrique orienté. On dit que $s$ *est l'abscisse curviligne de l'arc orienté.*

$ds = \|\gamma'(t)\|\,dt$ est appelé *élément de longueur.*

## 4.3.2 Aire d'un compact d'une nappe.

**Nappe géométrique de classe $C^1$.**

Une *nappe paramétrée de classe* $C^1$ est un couple $(U,f)$, où $U$ est un ouvert *connexe* (*i.e.* qui n'est pas réunion de deux ouverts disjoints non vides) de $\mathbb{R}^2$ et $f : (u_1,u_2) \mapsto f(u_1,u_2)$ une application de classe $C^1$ (*i.e.* continue dans $U$ et ayant des dérivées partielles d'ordre 1 continues dans $U$) de $U$ dans $\mathbb{R}^3$. Deux nappes paramétrées de classe $C^1$ $(U,f)$ et $(V,g)$ sont dites *équivalentes* s'il existe un $C^1$-difféomorphisme $\varphi : U \to V$, (*i.e.* une application $\varphi : (u_1,u_2) \mapsto (\varphi_1(u_1,u_2), \varphi_2(u_1,u_2))$ de $U$ sur $V$ qui soit bijective, continue et à dérivées partielles d'ordre 1 continues, et telle que le *Jacobien*

$$J(\varphi) = \begin{vmatrix} \frac{\partial \varphi_1}{\partial u_1} & \frac{\partial \varphi_1}{\partial u_2} \\ \frac{\partial \varphi_2}{\partial u_1} & \frac{\partial \varphi_2}{\partial u_2} \end{vmatrix}$$

ne s'annule pas dans $U$) vérifiant la relation

$$f = g \circ \varphi \tag{4.21}$$

On définit ainsi une relation d'équivalence sur l'ensemble des nappes paramétrées de classe $C^1$. Une classe d'équivalence est appelée *nappe géométrique de classe* $C^1$. Etant donnée une nappe géométrique représentée par la nappe paramétrée $(U, f)$, toute autre nappe paramétrée équivalente est encore appelée *paramétrage admissible* de la nappe. Tous les paramétrages admissibles ont la même image, appelée support de la nappe.

La nappe $(U, f)$ est dite *injective* si l'application $f$ est injective.

### Aire d'un morceau de nappe.

On considère une nappe paramétrée $(U, f)$ et un morceau $\Sigma$ obtenu en restreignant le paramétrage à un sous-ensemble $S$ de $U$ (*i.e.* $\Sigma = f(S)$) avec soit $S$ compact, soit $S$ ouvert d'adhérence compacte. L'aire de $\Sigma$ est définie par

$$\mathfrak{a}(\Sigma) = \iint_S \left\| \frac{\partial f}{\partial u_1} \wedge \frac{\partial f}{\partial u_2} \right\| du_1 du_2.$$

Cela ne dépend pas du paramétrage admissible utilisé pour la nappe géométrique. En effet, si $(V, g)$ est un autre paramétrage admissible, il existe un $C^1$-difféomorphisme $\varphi : (u_1, u_2) \mapsto \varphi(u_1, u_2) = (\varphi_1(u_1, u_2), \varphi_2(u_1, u_2))$ de $U$ sur $V$ tel que $f = g \circ \varphi$. On a alors en notant $\frac{\partial g}{\partial v_1}$ et $\frac{\partial g}{\partial v_2}$ les dérivées partielles de $g(v_1, v_2)$ pour $(v_1, v_2) \in V$

$$\begin{aligned} \frac{\partial f}{\partial u_1} &= \frac{\partial g}{\partial v_1}(\varphi(u_1, u_2)) \frac{\partial \varphi_1}{\partial u_1} + \frac{\partial g}{\partial v_2}(\varphi(u_1, u_2)) \frac{\partial \varphi_2}{\partial u_1} \\ \frac{\partial f}{\partial u_2} &= \frac{\partial g}{\partial v_1}(\varphi(u_1, u_2)) \frac{\partial \varphi_1}{\partial u_2} + \frac{\partial g}{\partial v_2}(\varphi(u_1, u_2)) \frac{\partial \varphi_2}{\partial u_2} \end{aligned}$$

donc

$$\frac{\partial f}{\partial u_1} \wedge \frac{\partial f}{\partial u_2} = \frac{\partial g}{\partial v_1}(\varphi(u_1, u_2)) \wedge \frac{\partial g}{\partial v_2}(\varphi(u_1, u_2)) \; J(\varphi).$$

On a ainsi

$$\begin{aligned} \iint_S \left\| \frac{\partial f}{\partial u_1} \wedge \frac{\partial f}{\partial u_2} \right\| du_1 du_2 &= \iint_S \left\| \frac{\partial g}{\partial v_1}(\varphi(u_1, u_2)) \wedge \frac{\partial g}{\partial v_2}(\varphi(u_1, u_2)) \right\| |J(\varphi)| \, du_1 du_2 \\ &= \iint_{\varphi(S)} \left\| \frac{\partial g}{\partial v_1} \wedge \frac{\partial g}{\partial v_2} \right\| \; dv_1 dv_2 \end{aligned}$$

d'après la formule de changement de variables dans les intégrales doubles.

$d\sigma = \left\| \frac{\partial f}{\partial u_1} \wedge \frac{\partial f}{\partial u_2} \right\| du_1 du_2$ est appelé *élément d'aire.*

**Exemple.** La nappe paramétrée $(U, f)$ définie par $U = \mathbb{R}^2$ et

$$f(u_1, u_2) = \begin{pmatrix} R \sin u_1 \cos u_2 \\ R \sin u_1 \sin u_2 \\ R \cos u_1 \end{pmatrix}$$

avec $R > 0$ a pour support la sphère de rayon $R$. Cette nappe n'est pas injective. Si l'on restreint le paramétrage à l'ouvert

$$\{(u_1, u_2) \in ]0, \pi[ \times ]0, 2\pi[\},$$

on obtient une nappe injective dont le support est la sphère privée du "demi-grand cercle" $\Sigma$ défini par $u_1 \in [0, \pi], u_2 = 0$. L'élément d'aire est $d\sigma = R^2 \sin u_1 \, du_1 du_2$, et $\Sigma$ est d'aire nulle. L'aire $\mathfrak{a}$ de la sphère est donc l'aire de la nappe injective :

$$\mathfrak{a} = R^2 \int_0^{\pi} \sin u_1 \, du_1 \int_0^{2\pi} du_2 = 4\pi R^2.$$

### 4.3.3 Centre de gravité d'un solide.

**Solide.**

Soit $\Sigma$ une partie compacte de $\mathbb{R}^3$ portant une répartition de masses. Nous dirons que $\Sigma$ est un *solide*. Nous nous limitons aux cas suivants :

1. $\Sigma$ est fini : $\Sigma$ est formé d'un ensemble fini de points $A_1, \dots, A_n$ affectés des masses $m_1, \dots, m_n$, $\sum_{i=1}^n m_i \neq 0$.

2. $\Sigma$ est continu et de l'un des types suivants :

(i) $\Sigma$ est *linéique*, *i.e.* est le support d'un arc paramétré $([a, b], \gamma)$ de classe $C^1$ avec des masses réparties suivant une *densité linéique* de masses qui est une fonction continue $\varrho : [a, b] \to \mathbb{R}$,

(ii) $\Sigma$ est *surfacique*, *i.e.* est le support d'un morceau compact d'une nappe paramétrée $(U, f)$ de classe $C^1$ : $\Sigma = f(K)$, $K$ compact de $U$, avec des masses réparties suivant une *densité surfacique* de masses qui est une fonction continue $\varrho : K \to \mathbb{R}$,

(iii) $\Sigma$ est *volumique*, *i.e.* est l'adhérence d'un ouvert borné connexe de $\mathbb{R}^3$, avec des masses réparties suivant une *densité volumique* de masses qui est une fonction continue $\varrho : \Sigma \to \mathbb{R}$,

Si $F$ est une fonction continue sur $\Sigma$, nous définirons le symbole $\int_\Sigma F(A) \, dm(A)$ (ou simplement $\int_\Sigma F \, dm$ s'il n'y a pas de confusion) comme étant :

- $\sum_{i=1}^n m_i F(A_i)$ si $\Sigma$ est fini;
- $\int_a^b F(\gamma(t)) \, \varrho(t) \, \|\gamma'(t)\| \, dt$ si $\Sigma$ est linéique;
- $\iint_K F(f(u_1, u_2)) \, \varrho(u_1, u_2) \, \|\frac{\partial f}{\partial u_1} \wedge \frac{\partial f}{\partial u_2}\| \, du_1 du_2$ si $\Sigma$ est surfacique;
- $\iiint_\Sigma F(x, y, z) \, \varrho(x, y, z) \, dx \, dy \, dz$ si $\Sigma$ est volumique.

La masse $m(\Sigma)$ est définie comme étant $m(\Sigma) = \int_\Sigma 1 \, dm(A)$. On la suppose $\neq 0$. $\Sigma$ est dit *homogène* si la densité (linéique, surfacique ou volumique) est une constante.

**Définition du centre de gravité.**

**Proposition 4. 8.** *Soit $O$ un point de $\mathbb{R}^3$ pris comme origine.*
*(i) Soit $G$ le point de $\mathbb{R}^3$ tel que*

$$\mathbf{OG} = \frac{1}{m(\Sigma)} \int_\Sigma \mathbf{OA} \, dm(A) . \tag{4.22}$$

*$G$ ne dépend pas du point $O$ utilisé comme origine.*
*(ii) $G$ est le seul point de $\mathbb{R}^3$ tel que*

$$\int_\Sigma \mathbf{GA} \, dm(A) = 0. \tag{4.23}$$

**Démonstration.**

(i) Si $O'$ est un autre point pris comme origine et si $G'$ est le point défini par $\mathbf{O'G'} = \frac{1}{m(\Sigma)} \int_\Sigma \mathbf{O'A}\, dm(A)$, on aura

$$\begin{aligned}
\mathbf{O'G'} &= \frac{1}{m(\Sigma)} \int_\Sigma (\mathbf{O'O} + \mathbf{OA})\, dm(A) \\
&= \frac{1}{m(\Sigma)} \int_\Sigma \mathbf{O'O}\, dm(A) + \frac{1}{m(\Sigma)} \int_\Sigma \mathbf{OA}\, dm(A) \\
&= \mathbf{O'O} + \mathbf{OG}
\end{aligned}$$

donc $G = G'$.

(ii) Il suffit de noter que pour un point $M \in \mathbb{R}^3$ quelconque, on a

$$\begin{aligned}
\int_\Sigma \mathbf{MA}\, dm(A) &= \int_\Sigma (\mathbf{OA} - \mathbf{OM})\, dm(A) \\
&= \int_\Sigma \mathbf{OA}\, dm(A) - m(\Sigma)\, \mathbf{OM} \\
&= m(\Sigma)(\mathbf{OG} - \mathbf{OM}) \\
&= m(\Sigma)\, \mathbf{MG}.
\end{aligned}$$

□

**Définition 4. 12.** *Le point $G$ défini par (4.22) est appelé le centre de gravité du solide $\Sigma$.*

Lorsque $\Sigma$ est fini, on retrouve la définition du barycentre.

**Exemple : centre de gravité d'une plaque plane homogène triangulaire.**

Soit une plaque plane homogène triangulaire $ABC$. Considérons la nappe injective $(U, f)$ avec $U = \mathbb{R}^2$ et $f(u_1, u_2) = \mathbf{OA} + u_1\mathbf{AB} + u_2\mathbf{AC}$ ($O$ est l'origine canonique de $\mathbb{R}^3$). L'élément d'aire est $d\sigma = \|\mathbf{AB} \wedge \mathbf{AC}\|\, du_1 du_2$. La plaque peut être considérée comme le morceau compact $\Sigma$ de la nappe injective $(U, f)$ correspondant au sous-ensemble compact

$$S = \{(u_1, u_2) \in [0,1] \times [0,1]\,;\, 0 \leqslant u_1 + u_2 \leqslant 1\} \subset U\,.$$

La densité surfacique $\varrho$ étant constante, le centre de gravité $G$ est défini (en prenant comme nouvelle origine $A$) par

$$\mathbf{AG} = \frac{1}{\iint_S d\sigma} \iint_S (u_1\mathbf{AB} + u_2\mathbf{AC})\, d\sigma.$$

Or en notant pour simplifier $a = \|\mathbf{AB} \wedge \mathbf{AC}\|$, on obtient :

$$\begin{aligned}
\iint_S d\sigma &= a \int_0^1 du_1 \int_0^{1-u_1} du_2 \\
&= a \int_0^1 (1 - u_1)\, du_1 \\
&= a \left[ -\frac{(1-u_1)^2}{2} \right]_0^1 \\
&= \frac{a}{2}
\end{aligned}$$

et

$$\iint_S (u_1\mathbf{AB} + u_2\mathbf{AC})\, d\sigma$$

$$\begin{aligned}
&= a\int_0^1 du_1 \int_0^{1-u_1} (u_1\mathbf{AB} + u_2\mathbf{AC})\, du_2 \\
&= a\int_0^1 \left(u_1(1-u_1)\mathbf{AB} + \frac{(1-u_1)^2}{2}\mathbf{AC}\right) du_1 \\
&= a\left(\left[\frac{u_1^2}{2} - \frac{u_1^3}{3}\right]_0^1 \mathbf{AB} + \left[-\frac{(1-u_1)^3}{6}\right]_0^1 \mathbf{AC}\right) \\
&= \frac{a}{6}(\mathbf{AB} + \mathbf{AC}).
\end{aligned}$$

D'où

$$\mathbf{AG} = \frac{1}{3}(\mathbf{AB} + \mathbf{AC}) = \frac{2}{3}\mathbf{AH}$$

puisque $\frac{1}{2}(\mathbf{AB} + \mathbf{AC}) = \mathbf{AH}$ en notant $H$ le milieu du segment $[BC]$. Ainsi $G$ est situé au $\frac{2}{3}$ de la médiane et est donc l'isobarycentre des points $A, B, C$.

Les propriétés du centre de gravité sont analogues à celles du barycentre, les démonstrations consistant essentiellement à remplacer $\sum_{i=1}^n \lambda_i OA_i$ par $\int_\Sigma OA\, dm$. En particulier l'analogue de la propriété d'associativité des barycentres (Th.4.2) est :

**Théorème 4. 5.** *Soit $\Sigma$ une partie compacte de $\mathbb{R}^3$ portant une répartition de masses. On suppose que $\Sigma$ est réunion de deux compacts $\Sigma_1$ et $\Sigma_2$ tels que $m(\Sigma_1 \cap \Sigma_2) = 0$. Si $G_i$ ($i = 1, 2$) désigne le centre de gravité de $\Sigma_i$, le centre de gravité de $\Sigma$ est le barycentre de $G_1$ et $G_2$ affectés des masses respectives $m(\Sigma_1)$ et $m(\Sigma_2)$.*

## 4.4 Exercices.

**Exercice 4.1.**

(i) Soit $P$ un polynôme à coefficients complexes. Montrer que l'ensemble des zéros du polynôme dérivé $P'$ est contenu dans l'enveloppe convexe de l'ensemble des zéros de $P$.

(ii) En déduire que le polynôme $1 + X + aX^n$, où $n \in \mathbb{N}, n \geqslant 2, a \in \mathbb{C}$, possède toujours au moins un zéro de module $\leqslant 2$.

**Indication.**

(i) Introduire les zéros distincts $\alpha_1, \ldots, \alpha_r$ de $P$ avec leurs multiplicités $k_1, \ldots, k_r$, et considérer $\frac{P'}{P}$.

**Exercice 4.2.**

Soit $E = \mathbb{R}_n[X]$ l'espace vectoriel de polynômes de degré $\leqslant n$ à coefficients réels, et $Y = \{P \in E;\ P(0) = 1,\ P(1) = 0\}$. Montrer que $Y$ est un sous-espace affine de $E$ et donner une base de sa direction.

**Exercice 4.3.**

Soient $A, B, C$ 3 points non alignés du plan affine. Quel est le lieu des points $D$ du plan ayant la propriété suivante : si $H$ (resp. $K$) désigne l'isobarycentre des 3

points $A, B, D$ (resp. $A, C, D$), et $G$ désigne l'isobarycentre des 4 points $A, B, C, D$, alors $G$ est sur la droite passant par $H$ et $K$?

**Indication.**

Prenons comme repère affine $(A, (\mathbf{AB}, \mathbf{AC}))$. $G$ appartient à la droite $HK$ si et seulement si il existe $\lambda \in \mathbb{R}$ tel que $\mathbf{AG} = \lambda\mathbf{AH} + (1-\lambda)\mathbf{AK}$. Si l'on pose $\mathbf{AD} = x\mathbf{AB} + y\mathbf{AC}$, cela s'écrit :

$$\frac{1}{4}(\mathbf{AB} + \mathbf{AC} + \mathbf{AD}) = \frac{\lambda}{3}(\mathbf{AB} + \mathbf{AD}) + \frac{1-\lambda}{3}(\mathbf{AC} + \mathbf{AD})$$

*i.e.*

$$\frac{1}{4}((1+x)\mathbf{AB} + (1+y)\mathbf{AC}) = \frac{1}{3}((x+\lambda)\mathbf{AB} + (y+1-\lambda)\mathbf{AC})$$

ou encore

$$\begin{cases} \lambda = \frac{3}{4} - \frac{x}{4} \\ \lambda = \frac{1}{4} + \frac{y}{4}. \end{cases}$$

Le lieu de $D$ est donc défini par $x + y = 2$. On a

$$\mathbf{AD} = x\mathbf{AB} + (2-x)\mathbf{AC} = 2\mathbf{AC} - x\mathbf{BC},$$

donc le lieu est la droite de vecteur directeur $\mathbf{BC}$ passant par le point $M$ tel que $\mathbf{AM} = 2\mathbf{AC}$. Dans le cas particulier où $A, B, C$ sont 3 sommets d'un rectangle, le lieu de $D$ est la parallèle à la diagonale $BC$ passant par le quatrième sommet du rectangle.

**Exercice 4.4.**

Dans l'espace affine euclidien $\mathbb{R}^n$, soit $B = \{M\,;\, \|\mathbf{OM}\| \leqslant 1\}$ la boule unité fermée de centre $O$. Montrer que l'ensemble des points extrêmaux de $B$ est la sphère $S = \{M\,;\, \|\mathbf{OM}\| = 1\}$ de centre $O$.

**Indication.**

Montrons d'abord que tout point de $S$ est un point extrêmal de $B$. Soit $M \in S$. Si $M$ n'était pas un point extrêmal de $B$, il existerait deux points distincts $P, Q \in B$ et $\lambda \in ]0,1[$ tels que $\mathbf{OM} = \lambda\mathbf{OP} + (1-\lambda)\mathbf{OQ}$. On aurait alors

$$\|\mathbf{OM}\|^2 = \lambda^2\|\mathbf{OP}\|^2 + (1-\lambda)^2\|\mathbf{OQ}\|^2 + 2\lambda(1-\lambda)(\mathbf{OP}|\mathbf{OQ}).$$

Or $\|\mathbf{OP}\| \leqslant 1$, $\|\mathbf{OQ}\| \leqslant 1$, et $(\mathbf{OP}|\mathbf{OQ}) \leqslant \|\mathbf{OP}\|\|\mathbf{OQ}\| \leqslant 1$ (inégalité de Cauchy-Schwarz). Si l'une de ces 3 inégalités était stricte, on aurait

$$\|\mathbf{OM}\|^2 < \lambda^2 + (1-\lambda)^2 + 2\lambda(1-\lambda) = (\lambda + (1-\lambda))^2 = 1,$$

en contradiction avec $\|\mathbf{OM}\|^2 = 1$. Donc $\|\mathbf{OP}\| = 1$, $\|\mathbf{OQ}\| = 1$ et $(\mathbf{OP}|\mathbf{OQ}) = 1$. Mais alors

$$\|\mathbf{OQ} - \mathbf{OP}\|^2 = \|\mathbf{OP}\|^2 + \|\mathbf{OQ}\|^2 - 2(\mathbf{OP}|\mathbf{OQ}) = 0,$$

ce qui est absurde puisque $P \neq Q$. Donc $M$ est un point extrêmal de $B$. Cela prouve que $S \subset Extr(B)$. Par ailleurs, si $M$ est un point de $B \setminus S$, on a $r = \|\mathbf{OM}\| < 1$, et les deux points $P, Q$ de $B$ définis par $\mathbf{OP} = \left(1 + \frac{1-r}{2r}\right)\mathbf{OM}$ et $\mathbf{OQ} = \left(1 - \frac{1-r}{2r}\right)\mathbf{OM}$ ont pour milieu $M$, donc $M$ n'est pas un point extrêmal de $B$. Cela prouve que $Extr(B) \subset S$. D'où finalement $Extr(B) = S$.

# Chapitre 5

# Groupes de symétries.

## 5.1 Compléments sur le groupe symétrique $\mathcal{S}_n$.

### 5.1.1 Décomposition en cycles disjoints.

**Théorème 5. 1.** *Toute permutation $\sigma \in \mathcal{S}_n$ se décompose en un produit de cycles disjoints. Cette décomposition est unique (à l'ordre près des cycles, car des cycles disjoints commutent entre eux).*

**Démonstration.**

*Existence.* Soit $G = \langle\sigma\rangle = \{I, \sigma, \sigma^2, \dots, \sigma^{k-1}\}$ le sous-groupe de $\mathcal{S}_n$ engendré par $\sigma$, $k$ désignant l'ordre de $\sigma$ dans $\mathcal{S}_n$. L'action naturelle de $\mathcal{S}_n$ sur $X = \{1, \dots, n\}$ donne par restriction à $G$ une action de $G$ sur $X$. L'orbite d'un élément $i \in X$ sous l'action de $G$ est

$$\mathcal{O}(i) = \{i, \sigma(i), \dots, \sigma^{k-1}(i)\}.$$

Mais il faut noter que les éléments $i, \sigma(i), \dots, \sigma^{k-1}(i)$ ne sont pas nécessairement deux-à-deux distincts. Soit $p_i$ le plus petit entier $p > 0$ tel que $\sigma^p(i) = i$. Alors les éléments $i, \sigma(i), \dots, \sigma^{p_i-1}(i)$ sont deux-à-deux distincts, donc l'orbite $\mathcal{O}(i)$ est formé de ces $p_i$ éléments :

$$\mathcal{O}(i) = \{i, \sigma(i), \dots, \sigma^{p_i-1}(i)\}.$$

La restriction de $\sigma$ à l'orbite $\mathcal{O}(i)$ est le $p_i$-cycle $c_i = (i, \sigma(i), \dots, \sigma^{p_i-1}(i))$.

Maintenant, $X$ est la réunion disjointe des différentes orbites sous l'action de $G$ :

$$\begin{aligned} X &= \mathcal{O}(i_1) \cup \dots \cup \mathcal{O}(i_s) \\ &= \{i_1, \sigma(i_1), \dots, \sigma^{p_{i_1}-1}(i_1)\} \cup \dots \cup \{i_s, \sigma(i_s), \dots, \sigma^{p_{i_s}-1}(i_s)\} \end{aligned}$$

Il en résulte que $\sigma$ est le produit des différents cycles obtenus par restriction à chaque orbite :

$$\begin{aligned} \sigma &= (i_1, \sigma(i_1), \dots, \sigma^{p_{i_1}-1}(i_1)) \dots (i_s, \sigma(i_s), \dots, \sigma^{p_{i_s}-1}(i_s)) \\ &= c_{i_1} \dots c_{i_s}. \end{aligned}$$

On notera que la décomposition ci-dessus peut comporter des 1-cycles, correspondant à des orbites triviales (*i.e.* réduites à un point).

*Unicité.* Soit $\sigma = \tau_1 \tau_2 \dots \tau_r$ une décomposition quelconque de $\sigma$ en cycles disjoints.

Dans cette décomposition, les éventuels 1-cycles sont écrits. Pour $1 \leqslant j \leqslant r$, $\tau_j$ est de la forme

$$\tau_j = (a_j, \sigma(a_j), \dots, \sigma^{q_j-1}(a_j)),$$

donc

$$\mathcal{O}(a_j) = \{a_j, \sigma(a_j), \dots, \sigma^{q_j-1}(a_j)\}$$

est l'orbite de $a_j$ sous $G$. La décomposition de $X$ en orbites sous $G$ est donc :

$$X = \mathcal{O}(a_1) \cup \mathcal{O}(a_2) \cup \dots \cup \mathcal{O}(a_r).$$

On déduit de l'unicité de la décomposition en orbites que $r = s$ et que, à l'ordre près,

$$\mathcal{O}(a_1) = \mathcal{O}(i_1),\ \mathcal{O}(a_2) = \mathcal{O}(i_2), \dots, \mathcal{O}(a_s) = \mathcal{O}(i_s).$$

Cela implique $\tau_i = c_i \quad \forall i = 1, \dots, s$. □

**Exemple.** Dans la pratique la décomposition en cycles disjoints s'obtient comme dans la démonstration du théorème. Prenons comme exemple la permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 5 & 8 & 6 & 4 & 10 & 1 & 3 & 2 & 7 & 9 \end{pmatrix} \in \mathcal{S}_{10}.$$

On commence avec 1 : 1 donne 5, qui donne 10 , qui donne 9, qui donne 7, qui donne 3, qui donne 6, qui redonne 1, d'où le 7-cycle $(1, 5, 10, 9, 7, 3, 6)$ correspondant à l'orbite de 1. On recommence ensuite avec un nombre différent de ceux obtenus, par exemple 2, et l'on obtient de même le 2-cycle (*i.e.* la transposition) $(2, 8)$. On termine avec 4 qui donne le 1-cycle (*i.e.* point fixe) $(4)$. La décomposition de $\sigma$ en cycles disjoints est donc

$$(1, 5, 10, 9, 7, 3, 6)(2, 8)(4)$$

On n'écrit en général pas les 1-cycles, de sorte que la décomposition est simplement :

$$(1, 5, 10, 9, 7, 3, 6)(2, 8).$$

**Corollaire.** *Toute permutation $\sigma \in \mathcal{S}_n$ se décompose en un produit de transpositions. Ces transpositions ne sont pas nécessairement disjointes, et en conséquence, cette décomposition n'est pas nécessairement unique.*

**Démonstration.**

Il suffit de remarquer que si $\sigma$ est un $p$-cycle quelconque $(a_1, a_2, \dots, a_p)$, avec une suite $1 \leqslant a_1, a_2, \dots, a_p \leqslant n$ pas nécessairement croissante, on a

$$(a_1, a_2, \dots, a_p) = (a_1, a_p)(a_1, a_{p-1}) \dots (a_1, a_3)(a_1, a_2). \tag{5.1}$$

Pour la non-unicité, il suffit d'observer que l'on a aussi la décomposition

$$(a_1, a_2, \dots, a_p) = (a_1, a_2)(a_2, a_3) \dots (a_{p-2}, a_{p-1})(a_{p-1}, a_p). \tag{5.2}$$

□

### 5.1.2 Homomorphisme signature.

**Définition 5. 1.** *Soit* $\sigma \in \mathcal{S}_n$. *On dit qu'un couple* $(i,j)$ *tel que* $1 \leqslant i < j \leqslant n$ *présente une inversion si* $\sigma(i) > \sigma(j)$*; on appelle nombre d'inversions de* $\sigma$ *l'entier* $N = \operatorname{card}\{(i,j);\ 1 \leqslant i < j \leqslant n\ ;\ \sigma(i) > \sigma(j)\}$. *On appelle signature de* $\sigma$ *le nombre* $\varepsilon(\sigma) = (-1)^N$.

**Proposition 5. 1.** *(i)* $\varepsilon(\sigma) = \prod_{1 \leqslant i<j \leqslant n} \frac{\sigma(j)-\sigma(i)}{j-i}$.
*(ii) L'application* $\varepsilon : \mathcal{S}_n \longrightarrow \{-1,+1\}$ *est un homomorphisme du groupe* $\mathcal{S}_n$ *dans le groupe multiplicatif* $\{-1,+1\}$.

**Démonstration.**
(i)

$$\prod_{1\leqslant i<j\leqslant n} \frac{\sigma(j)-\sigma(i)}{j-i} = \frac{\prod_{1\leqslant i<j\leqslant n}(\sigma(j)-\sigma(i))}{\prod_{1\leqslant \ell<m\leqslant n}(m-\ell)}.$$

Considérons le numérateur

$$\prod_{1\leqslant i<j\leqslant n} (\sigma(j)-\sigma(i)).$$

On a pour $1 \leqslant i < j \leqslant n$

$$\sigma(j)-\sigma(i) = \pm(\sup(\sigma(i),\sigma(j)) - \inf(\sigma(i),\sigma(j))),$$

avec le signe + si $i$ et $j$ ne présentent pas une inversion et le signe − sinon. Donc, en notant $N$ le nombre d'inversions de $\sigma$,

$$\begin{aligned}\prod_{1\leqslant i<j\leqslant n} (\sigma(j)-\sigma(i)) &= (-1)^N \prod_{1\leqslant i<j\leqslant n} (\sup(\sigma(i),\sigma(j)) - \inf(\sigma(i),\sigma(j))) \\ &= \varepsilon_\sigma \prod_{1\leqslant i<j\leqslant n} (\sup(\sigma(i),\sigma(j)) - \inf(\sigma(i),\sigma(j))).\end{aligned}$$

Or l'application

$$(i,j) \mapsto (\inf(\sigma(i),\sigma(j)), \sup(\sigma(i),\sigma(j))) \tag{5.3}$$

est une bijection de $\{(i,j);\ 1 \leqslant i < j \leqslant n\}$ sur lui-même, *i.e.* pour tout couple $(\ell, m)$ tel que $1 \leqslant \ell < m \leqslant n$ il existe un couple unique $(i,j)$ tel que $1 \leqslant i < j \leqslant n$ pour lequel

$$\ell = \inf(\sigma(i),\sigma(j)) \ \text{ et } \ m = \sup(\sigma(i),\sigma(j)).$$

En effet, on doit avoir $\{\ell, m\} = \{\sigma(i),\sigma(j)\}$ donc $\sigma(i) = \ell$ et $\sigma(j) = m$ ou $\sigma(i) = m$ et $\sigma(j) = \ell$. Il n'y a donc que deux couples susceptibles de convenir : $(\sigma^{-1}(\ell), \sigma^{-1}(m))$ et $(\sigma^{-1}(m), \sigma^{-1}(\ell))$. Il y a bien un et un seul de ces deux couples qui est dans l'ordre croissant.

On a donc

$$\prod_{1\leqslant i<j\leqslant n} (\sup(\sigma(i),\sigma(j)) - \inf(\sigma(i),\sigma(j))) = \prod_{1\leqslant \ell<m\leqslant n} (m-\ell)$$

et alors

$$\prod_{1\leqslant i<j\leqslant n} \frac{\sigma(j)-\sigma(i)}{j-i} = \varepsilon(\sigma)\frac{\prod_{1\leqslant i<j\leqslant n}(\sup(\sigma(i),\sigma(j)) - \inf(\sigma(i),\sigma(j)))}{\prod_{1\leqslant \ell<m\leqslant n}(m-\ell)} = \varepsilon(\sigma).$$

(ii) On a pour $\tau, \sigma \in \mathcal{S}_n$ :

$$\begin{aligned}\varepsilon(\tau\sigma) &= \prod_{1\leqslant i<j\leqslant n} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{j-i} \\ &= \prod_{1\leqslant i<j\leqslant n} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j)-\sigma(i)} \prod_{1\leqslant i<j\leqslant n} \frac{\sigma(j)-\sigma(i)}{j-i}.\end{aligned}$$

Or on a

$$\prod_{1\leqslant i<j\leqslant n} \frac{\sigma(j)-\sigma(i)}{j-i} = \varepsilon(\sigma),$$

et d'autre part

$$\begin{aligned}\prod_{1\leqslant i<j\leqslant n} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j)-\sigma(i)} &= \prod_{1\leqslant i<j\leqslant n} \frac{\tau(\sup(\sigma(i),\sigma(j))) - \tau(\inf(\sigma(i),\sigma(j)))}{\sup(\sigma(i),\sigma(j)) - \inf(\sigma(i),\sigma(j))} \\ &= \prod_{1\leqslant \ell<m\leqslant n} \frac{\tau(m)-\tau(\ell)}{m-\ell} \\ &\quad \text{(d'après la bijectivité de l'application (5.3))} \\ &= \varepsilon(\tau).\end{aligned}$$

Donc

$$\varepsilon(\tau\sigma) = \varepsilon(\tau)\,\varepsilon(\sigma).$$

□

### 5.1.3 Exemples.

(i) *La signature d'une transposition est* $-1$.

Si $\sigma = (1,2)$, le seul couple $(i,j)\,(1 \leqslant i < j \leqslant n)$ présentant une inversion est le couple $(1,2)$. Le nombre d'inversions de $\sigma$ est donc 1 et la signature $-1$.

Si $\sigma = (i,j)$ est une transposition quelconque $(1 \leqslant i < j \leqslant n)$, on a d'après la formule fondamentale (1.9)

$$\sigma = s(1,2)s^{-1}$$

avec $s = (1,i)(2,j)$, de sorte que $\varepsilon(\sigma) = \varepsilon(s)\varepsilon((1,2))\varepsilon(s)^{-1} = \varepsilon((1,2)) = -1$.

(ii) *La signature d'un $p$-cycle est* $(-1)^{p-1}$.

Si $\sigma$ est le $p$-cycle $(1,2,\ldots,p)$, on a

$$(1,2,\ldots,p) = (1,p)(1,p-1)\ldots(1,3)(1,2)$$

donc $\varepsilon(\sigma) = (-1)^{p-1}$.

Si $\sigma$ est un $p$-cycle quelconque $(a_1,a_2,\ldots,a_p)$, avec une suite $1 \leqslant a_1, a_2, \ldots, a_p \leqslant n$ pas nécessairement croissante, on a de même (formule (5.1))

$$(a_1,a_2,\ldots,a_p) = (a_1,a_p)(a_1,a_{p-1})\ldots(a_1,a_3)(a_1,a_2)$$

donc $\varepsilon(\sigma) = (-1)^{p-1}$.

**Théorème 5. 2.** *$\mathcal{A}_n = \{\sigma \in \mathcal{S}_n;\ \varepsilon(\sigma) = 1\}$ est un sous-groupe distingué de $\mathcal{S}_n$. Son cardinal est $\frac{n!}{2}$.*

**Démonstration.**

$\mathcal{A}_n$ est le noyau de l'homomorphisme signature, donc c'est un sous-groupe distingué de $\mathcal{S}_n$. Par décomposition canonique de l'homomorphisme signature, on obtient un isomorphisme du groupe quotient $\mathcal{S}_n/\mathcal{A}_n$ sur le groupe multiplicatif $\{-1,1\}$, donc $\mathrm{card}(\mathcal{S}_n/\mathcal{A}_n) = 2$ et $\mathrm{card}(\mathcal{A}_n) = \frac{n!}{2}$. □

**Définition 5. 2.** *Le groupe $\mathcal{A}_n$ s'appelle le groupe alterné d'ordre $n$.*

## 5.2 Groupe diédral $D_n$, $n \geqslant 3$.

Soit $P_n$ un polygone convexe régulier à $n$ côtés ($n \geqslant 3$) du plan affine euclidien orienté $E$, $O$ son centre et $A_0, \dots, A_{n-1}$ ses sommets.

Considérons un repère affine orthonormé direct $(O,(\mathbf{e_1},\mathbf{e_2}))$ de $E$, avec $\mathbf{OA_0} = \lambda\mathbf{e_1}$ et la base $(\mathbf{e_1},\mathbf{e_2})$ de l'espace vectoriel euclidien $\mathbb{R}^2$ de même orientation que la base $(\mathbf{OA_0},\mathbf{OA_1})$. On peut supposer $\lambda = 1$. Un point $M \in E$ est identifié au vecteur $\mathbf{OM} \in \mathbb{R}^2$.

On a

$$\mathbf{OA_k} = r^k(\mathbf{OA_0}) \quad \forall k,\ 0 \leqslant k \leqslant n-1$$

en notant $r$ la rotation d'angle $\frac{2\pi}{n}$ dont la matrice dans la base orthonormée $(\mathbf{e_1},\mathbf{e_2})$ de l'espace vectoriel $\mathbb{R}^2$ est

$$R(\frac{2\pi}{n}) = \begin{pmatrix} \cos\frac{2\pi}{n} & -\sin\frac{2\pi}{n} \\ \sin\frac{2\pi}{n} & \cos\frac{2\pi}{n} \end{pmatrix}.$$

On peut d'autre part identifier l'espace vectoriel $\mathbb{R}^2$ au plan complexe $\mathbb{C}$ considéré comme espace vectoriel réel : le vecteur $\mathbf{OM} = x\,\mathbf{e_1} + y\,\mathbf{e_2} \in \mathbb{R}^2$ est identifié au nombre complexe $z = x+iy$. L'endomorphisme $r$ de $\mathbb{R}^2$ s'identifie alors à l'application $z \mapsto e^{i\frac{2\pi}{n}} z$, *i.e.* la multiplication par $e^{i\frac{2\pi}{n}}$ dans l'espace vectoriel réel $\mathbb{C}$.

Avec ces notations, les sommets $A_0, \dots, A_{n-1}$ de $P_n$ sont tels que $\mathbf{OA_k} = z_k$ où $z_k = e^{\frac{2ik\pi}{n}}$, $0 \leqslant k < n$.

Notons que $O$ est l'isobarycentre des points $A_0, \dots, A_{n-1}$. En effet,

$$\sum_{k=0}^{n-1} \mathbf{OA_k} = \sum_{k=0}^{n-1} z_k = \sum_{k=0}^{n-1} e^{\frac{2ik\pi}{n}} = \sum_{k=0}^{n-1} (e^{i\frac{2\pi}{n}})^k = \frac{1-(e^{i\frac{2\pi}{n}})^n}{1-e^{i\frac{2\pi}{n}}} = 0\,.$$

Soit $f$ une isométrie de $E$. On sait que $f$ est une bijection affine (Th.4.1). Si $g$ désigne la partie linéaire de $f$, on a

$$f(M) = f(O) + g(\mathbf{OM}) \quad \forall M \in E.$$

Supposons maintenant que l'isométrie $f$ laisse invariant $P_n$, *i.e.* $f(P_n) = P_n$. D'après la Prop.4.7, $f$ laisse aussi invariant l'ensemble des points extrêmaux de $P_n$ : $f(Extr(P_n)) = Extr(P_n)$. L'ensemble des points extrêmaux de $P_n$ est l'ensemble des sommets de $P_n$ d'après l'exemple (ii) de la section 4.2.6. Donc les ensembles $\{f(A_0), \dots, f(A_{n-1})\}$ et $\{A_0, \dots, A_{n-1}\}$ sont égaux. Or $O$ est l'isobarycentre des points $A_0, \dots, A_{n-1}$ et $f$ est affine, donc $f(O)$ est l'isobarycentre des points $f(A_0), \dots, f(A_{n-1})$. Ces ensembles de points étant les mêmes, on a $f(O) = O$. Ainsi

$$f(M) = O + g(\mathbf{OM}) \quad \forall M \in E.$$

Quand on identifie un point $M \in E$ au vecteur $\mathbf{OM} \in \mathbb{R}^2$, $f$ est alors identifiée à sa partie linéaire $g$ qui est un endomorphisme isométrique de l'espace vectoriel $\mathbb{R}^2$, *i.e.* un élément de $O(\mathbb{R}^2)$.

**Théorème 5. 3.** *Dans l'espace vectoriel euclidien $\mathbb{R}^2$ identifié à l'espace vectoriel réel $\mathbb{C}$, soit $P_n$ $(n \geqslant 3)$ le polygone dont les sommets sont les $z_k = e^{\frac{2ik\pi}{n}}$, $0 \leqslant k < n$. Soit $D_n$ l'ensemble des endomorphismes isométriques de $\mathbb{R}^2$ qui laissent stable $P_n$.*
*(i) $D_n$ est un sous-groupe de $O(\mathbb{R}^2)$.*
*(ii) $D_n$ est engendré par deux éléments $r, s$ tels que*

$$r \text{ est d'ordre } n,\ s \text{ est d'ordre } 2,\ s \notin \langle r \rangle \text{ et } \quad srs = r^{-1}. \tag{5.4}$$

*(iii) $D_n$ est d'ordre $2n$ et en tant qu'ensemble*

$$D_n = \{Id, r, r^2, \dots, r^{n-1}, s, sr, sr^2, \dots, sr^{n-1}\}.$$

*(iv) Tout groupe engendré par deux éléments ayant les propriétés (5.4) est isomorphe à $D_n$.*

**Démonstration.**

(i) Il est clair que l'identité conserve $P_n$, que la composée de deux isométries laissant invariant $P_n$ laisse encore invariant $P_n$, et que l'application réciproque d'une telle isométrie laisse invariant $P_n$. Donc $D_n$ est un sous-groupe de $O(\mathbb{R}^2)$.

(iii) Un élément $f \in O(\mathbb{R}^2)$ a pour déterminant $\pm 1$. Cherchons d'abord les éléments de $D_n$ de déterminant 1, *i.e.* les rotations de $\mathbb{R}^2$ conservant $P_n$. Les seules rotations qui conservent $P_n$ sont $Id, r, r^2, \dots, r^{n-1}$ en notant $r$ la rotation d'angle $\frac{2\pi}{n}$, *i.e.* la la multiplication par $e^{i\frac{2\pi}{n}}$. Cherchons maintenant les éléments de $D_n$ de déterminant $-1$. Soit $s$ la symétrie par rapport à "l'axe de $x$", *i.e.* la conjugaison complexe de $\mathbb{C}$ : $s(z) = \bar{z}\ \forall z \in \mathbb{C}$. On a $s(z_0) = z_0$ et pour $1 \leqslant k < n$, $s(z_k) = e^{-\frac{2ik\pi}{n}} = e^{\frac{2i(n-k)\pi}{n}} = z_{n-k}$ d'où $s \in D_n$. La matrice de $s$ dans la base canonique est $J = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, donc $\det s = -1$. Un élément $\sigma \in D_n$ est de déterminant $-1$ si et seulement si $s\sigma$ est un élément de $D_n$ de déterminant 1. Les éléments de $D_n$ de déterminant $-1$ sont donc les $\sigma = sr^k$, $0 \leqslant k < n$. Ainsi

$$D_n = \{Id, r, r^2, \dots, r^{n-1}, s, sr, sr^2, \dots, sr^{n-1}\}.$$

(ii) Il est clair que $D_n = \langle r, s \rangle$, que $r$ est d'ordre $n$, et $s$ d'ordre 2. $s \notin \langle r \rangle$ puisque $\det s = -1$. Calculons $srs$. Pour tout $z \in \mathbb{C}$ on a :

$$srs\,(z) = s(r(s(z))) = s(r(\bar{z})) = s(e^{\frac{2i\pi}{n}}\bar{z}) = \overline{e^{\frac{2i\pi}{n}}\bar{z}} = e^{-\frac{2i\pi}{n}}z = r^{-1}(z).$$

Donc $srs = s^{-1}$.

(iv) Soit $G$ un groupe engendré par deux éléments $a, b$ tels que $a$ soit d'ordre $n$, $b$ d'ordre 2, $b \notin \langle a \rangle$ et $bab = a^{-1}$. D'après le Lemme 1.5, tout élément $x \in G$ est de la forme $x = x_1 \dots x_p$ pour un certain $p \in \mathbb{N}^*$ avec pour tout $i$, $(1 \leqslant i \leqslant p)$, $x_i = a$ ou $x_i = a^{-1} = a^{n-1}$ ou $x_i = b = b^{-1}$. Comme $bab = a^{-1} = a^{n-1}$, on a $ab = ba^{n-1}$ et par conséquent $x$ s'écrit sous la forme $x = b^\varepsilon a^k$ avec $\varepsilon = 0$ ou 1 et $0 \leqslant k < n$. Mais

il est facile de vérifier que les éléments $e, a, \dots a^{n-1}, b, ba, \dots ba^{n-1}$ sont deux-à-deux distincts. Donc $G$ est d'ordre $2n$ et

$$G = \{e, a, \dots a^{n-1}, b, ba, \dots ba^{n-1}\}.$$

Considérons alors l'application $\varphi : G \to D_n$ définie par $\varphi(b^\varepsilon a^k) = s^\varepsilon r^k$. Par définition, cette application est une bijection. Vérifions que c'est un homomorphisme de groupes, *i.e.*

$$\varphi(xy) = \varphi(x)\varphi(y) \quad \forall x, y \in G.$$

Plusieurs cas sont à distinguer.
• $x = a^k$, $y = a^\ell$. Dans ce cas, $xy = a^{k+\ell} = a^i$ avec $k+\ell \equiv i \pmod n$ et $0 \leqslant i < n$. Alors $\varphi(xy) = r^i = r^k r^\ell = \varphi(x)\varphi(y)$.
• $x = ba^k$, $y = a^\ell$. Dans ce cas, $xy = ba^{k+\ell} = ba^i$ avec $k + \ell \equiv i \pmod n$ et $0 \leqslant i < n$. Alors $\varphi(xy) = sr^i = sr^k r^\ell = \varphi(x)\varphi(y)$.
• $x = a^k$, $y = ba^\ell$. Dans ce cas, $xy = a^k ba^\ell = ba^{-k+\ell} = sa^i$ avec $-k+\ell \equiv i \pmod n$ et $0 \leqslant i < n$. Alors $\varphi(xy) = sr^i = sr^{-k}r^\ell = r^k sr^\ell = \varphi(x)\varphi(y)$.
• $x = ba^k$, $y = ba^\ell$. Dans ce cas, $xy = ba^k ba^\ell = a^{-k+\ell} = a^i$ avec $-k+\ell \equiv i \pmod n$ et $0 \leqslant i < n$. Alors $\varphi(xy) = r^i = r^{-k}r^\ell = sr^k sr^\ell = \varphi(x)\varphi(y)$. □

**Définition 5. 3.** *Le groupe $D_n$ ($n \geqslant 3$) est appelé groupe diédral d'indice $n$.*

## 5.3 Groupe des isométries du cube et du tétraèdre.

### 5.3.1 Groupe des isométries du cube.

**Conservation du centre.**

Soit $C$ un cube de l'espace affine euclidien à 3 dimensions $E$ et $O$ son centre. On peut supposer que la longueur des côtés est 1. Le centre $O$ est aussi l'isobarycentre des 8 sommets. Soit $f$ une isométrie de $E$. On sait que $f$ est une bijection affine (Th.4.1). En particulier, si $g$ désigne la partie linéaire de $f$, on a

$$f(M) = f(O) + g(\mathbf{OM}) \quad \forall M \in E.$$

Supposons maintenant que l'isométrie $f$ laisse invariant $C$, *i.e.* $f(C) = C$. D'après la Prop.4.7, $f$ laisse aussi invariant l'ensemble des points extrêmaux de $C$ : $f(Extr(C)) = Extr(C)$. L'ensemble des points extrêmaux de $C$ est l'ensemble des sommets de $C$ d'après l'exemple (iv) de la section 4.2.6. L'ensemble des sommets de $C$ est donc stable par $f$. Or $f$ est affine, donc $f(O)$ est l'isobarycentre des images des sommets, *i.e.* $f(O) = O$. (Pour montrer que l'ensemble des sommets de $C$ est stable par $f$, on peut aussi remarquer que pour deux points $M, N$ quelconques du cube, on a $\|\mathbf{MN}\| \leqslant \sqrt{3}$, l'égalité n'ayant lieu que si $M$ et $N$ sont deux sommets opposés. Pour deux sommets opposés $M, N$ on aura $\|\mathbf{f(M)f(N)}\| = \|\mathbf{MN}\| = \sqrt{3}$, donc $f(M)$ et $f(N)$ sont encore deux sommets opposés de $C$ et ainsi l'ensemble des sommets de $C$ est stable par $f$).

On choisit comme origine $O$. Comme $f(O) = O$, $f$ peut être identifié à sa partie linéaire $g$ qui est un endomorphisme isométrique de l'espace vectoriel euclidien $\mathbb{R}^3$, *i.e.* un élément de $O(\mathbb{R}^3)$.

Comme usuel, on identifie $O(\mathbb{R}^3)$ et $O(3)$.

**Groupe des rotations du cube.**

Soit $A = R_{\mathbf{k}}(\theta) \in SO(3)$ une rotation différente de l'identité, d'angle $\theta$ et d'axe $\Delta = \mathbb{R}\mathbf{k}$. On suppose que $A$ laisse invariant le cube, *i.e.* $A(C) = C$. L'axe $\Delta$ passant par $O$, 3 cas sont possibles.

**1er cas :** $\Delta$ ne rencontre aucune arête. Dans ce cas, la droite $\Delta$ passe par les milieux de deux faces opposées. Si $r = R_{\mathbf{k}}(\frac{\pi}{2})$, $A$ est l'une des trois rotations $r, r^2, r^3$. Le sous-groupe de $SO(3)$ engendré par $r$ est d'ordre 4. On dit que $\Delta$ est un *axe quadruple*. Il y a donc 3 axes quadruples joignant les milieux de faces opposées. Chacun de ces axes quadruples donne 3 rotations différentes de l'identité laissant invariant le cube.

**2ème cas :** $\Delta$ passe par un sommet. Dans ce cas, la droite $\Delta$ passe par aussi par le sommet opposé, donc c'est une diagonale du cube. Si $s = R_{\mathbf{k}}(\frac{2\pi}{3})$, $A$ est l'une des deux rotations $s, s^2$. Le sous-groupe de $SO(3)$ engendré par $s$ est d'ordre 3. On dit que $\Delta$ est un *axe triple*. Il y a donc 4 axes triples : les 4 diagonales. Chacun de ces axes triples donne 2 rotations différentes de l'identité laissant invariant le cube.

**3ème cas :** $\Delta$ rencontre une arête, mais ne passe pas par un sommet. Dans ce cas, la droite $\Delta$ passe par le milieu de l'arête, et donc aussi par le milieu de l'arête opposée, et si $t = R_{\mathbf{k}}(\pi)$, $A = t$. Le sous-groupe de $SO(3)$ engendré par $t$ est d'ordre 2. On dit que $\Delta$ est un *axe double*. Il y a donc 6 axes doubles. Chacun de ces axes donne 1 rotation différente de l'identité laissant invariant le cube.

En résumé on a donc

| | | | |
|---|---|---|---|
| 3 | axes quadruples | qui donnent | $3 \cdot 3 = 9$ rotations $\neq Id$ |
| 4 | axes triples | qui donnent | $4 \cdot 2 = 8$ rotations $\neq Id$ |
| 6 | axes doubles | qui donnent | $6 \cdot 1 = 6$ rotations $\neq Id$ |
| | | | $Id = 1$ rotation |
| | | | **total = 24 rotations** |

**Théorème 5. 4.** *Soit* $\text{Rot}\,(C) = \{f \in SO(\mathbb{R}^3);\ f(C) = C\}$ *l'ensemble des rotations laissant invariant le cube* $C$. $\text{Rot}\,(C)$ *est un sous-groupe de* $SO(\mathbb{R}^3)$ *isomorphe au groupe symétrique* $\mathcal{S}_4$.

**Démonstration.**

Il est clair que l'identité conserve le cube $C$, que la composée de deux rotations laissant invariant $C$ laisse encore invariant $C$, et que la rotation inverse d'une telle rotation laisse invariant $C$. Donc $\text{Rot}\,(C)$ est un sous-groupe de $SO(\mathbb{R}^3)$. On a vu que $|\text{Rot}\,(C)| = 24$. Montrons que $\text{Rot}\,(C)$ est isomorphe à $\mathcal{S}_4$ dont le cardinal est aussi 24. Soit $f \in \text{Rot}\,(C)$. Notons $d_i$ $(1 \leqslant i \leqslant 4)$ les 4 diagonales de $C$. L'image par $f$ de $d_i$ est une diagonale $d_j$. Notons $j = \sigma_f(i)$. Alors l'application $\sigma_f : \{1,2,3,4\} \longmapsto \{1,2,3,4\}$ est une bijection, *i.e.* $\sigma_f \in \mathcal{S}_4$. Par définition, on a

$$\sigma_{fog} = \sigma_f \sigma_g \quad \forall f, g \in \text{Rot}\,(C)\,.$$

En effet,

$$d_{\sigma_{fog}(i)} = (f \circ g)(d_i) = f(g(d_i)) = f(d_{\sigma_g(i)}) = d_{\sigma_f(\sigma_g(i))}.$$

Donc l'application $\sigma : \text{Rot}\,(C) \longrightarrow \mathcal{S}_4$ définie par $f \mapsto \sigma_f$ est un homomorphisme de groupes. Montrons que $\text{Ker}\,\sigma = \{Id\}$. Soit $f \in \text{Ker}\,\sigma$, $f \neq Id$. On a $\sigma_f = Id_{\{1,2,3,4\}}$,

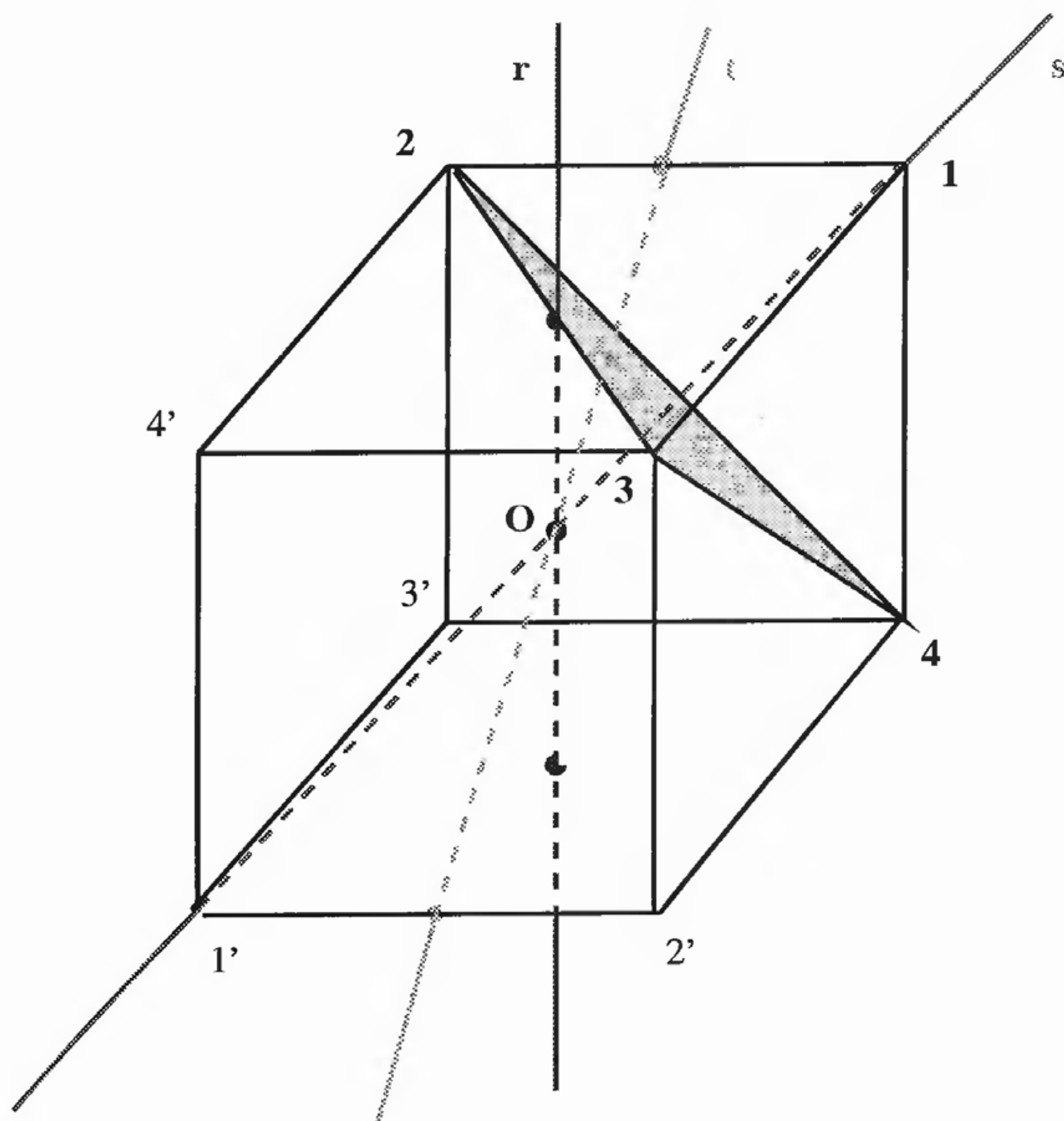



FIG. 5.1: *Les rotations du cube.*

*i.e.* $d_{\sigma_f(i)} = d_i \ \forall i = 1,2,3,4$. Cela signifie que la rotation $f$ laisse globalement invariante chacune des 4 diagonales du cube. Pour chaque $i\,(1 \leqslant i \leqslant 4)$, soit $\mathbf{u_i}$ un vecteur directeur normé de $d_i$. On a alors :

$$f(\mathbf{u_i}) = \varepsilon_i \mathbf{u_i} \quad \text{avec } \varepsilon_i = \pm 1 .$$

Considérons le triplet de vecteurs $(\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3})$. Cest une base de $\mathbb{R}^3$. La matrice $B$ de $f$ dans cette base est

$$B = \begin{pmatrix} \varepsilon_1 & 0 & 0 \\ 0 & \varepsilon_2 & 0 \\ 0 & 0 & \varepsilon_3 \end{pmatrix} .$$

Comme $\det B = \det f = 1$, on voit qu'il existe un indice $i_1 \in \{1,2,3\}$ unique tel que $\varepsilon_{i_1} = 1$, les deux autres indices $i_2, i_3$ restants étant tels que $\varepsilon_{i_2} = \varepsilon_{i_3} = -1$. On a $f(\mathbf{u_{i_1}}) = \mathbf{u_{i_1}}$ donc $\mathbf{u_{i_1}}$ appartient à l'axe de $f$. Considérons maintenant le triplet de vecteurs $(\mathbf{u_{i_2}}, \mathbf{u_{i_3}}, \mathbf{u_4})$. Comme $\varepsilon_{i_2} = \varepsilon_{i_3} = -1$, on a $\varepsilon_4 = 1$. Donc $\mathbf{u_4}$ appartient à l'axe de $f$. Ainsi l'axe de $f$ contiendrait $\{\mathbf{u_{i_1}}, \mathbf{u_4}\}$ et serait de dimension $\geqslant 2$ ce qui est absurde. Donc il n'existe pas d'élément $f \in \operatorname{Ker}\sigma$ différent de $Id$ et $\operatorname{Ker}\sigma = \{Id\}$. Cela prouve que l'homomorphisme $\sigma$ est injectif. Comme $\operatorname{Rot}(C)$ et $\mathcal{S}_4$ ont le même cardinal, $\sigma$ est un isomorphisme. □

**Exemple.** Avec les notations de la Figure 5.1

$$\begin{aligned} \sigma_r &= (1,2,4,3) \\ \sigma_s &= (1)(2,3,4) \\ \sigma_t &= (1,2)(3)(4). \end{aligned}$$

**Groupe des isométries du cube.**

**Théorème 5. 5.** *Soit* $\operatorname{Isom}(C) = \{f \in O(\mathbb{R}^3);\ f(C) = C\}$ *l'ensemble des isométries laissant invariant le cube* $C$. $\operatorname{Isom}(C)$ *est un sous-groupe de* $O(\mathbb{R}^3)$ *isomorphe au groupe produit direct*

$$\mathbb{Z}_2 \times \mathcal{S}_4.$$

**Démonstration.**

Il est immédiat que $\operatorname{Isom}(C)$ est un sous-groupe de $O(\mathbb{R}^3)$. Soit $h = -Id \in O(\mathbb{R}^3)$ la symétrie par rapport à l'origine $O$ : $h(\mathbf{OM}) = -\mathbf{OM}\ \forall M$. On a $h(C) = C$ et $\det h = -1$. Pour $f \in O(\mathbb{R}^3)$, $f \in \operatorname{Isom}(C)$ et $\det f = -1$ si et seulement si $g = h \circ f \in \operatorname{Rot}(C)$. Donc

$$\operatorname{Isom}(C) = \operatorname{Rot}(C) \cup (h \circ \operatorname{Rot}(C)) .$$

Toute $f \in \operatorname{Isom}(C)$ s'écrit donc sous la forme $f = h^\varepsilon \circ g$ avec $\varepsilon = 0$ ou $1$ et $g \in \operatorname{Rot}(C)$. Cette écriture est par ailleurs unique. Soit $\varphi : \operatorname{Isom}(C) \to \mathbb{Z}_2 \times \mathcal{S}_4$ définie par $\varphi(f) = ([\varepsilon], \sigma_g)$. Montrons que $\varphi$ est un homomorphisme de groupes. Pour $f = h^\varepsilon \circ g$ et $f' = h^{\varepsilon'} \circ g'$, on a

$$f \circ f' = h^\varepsilon \circ g \circ h^{\varepsilon'} \circ g' = h^{\varepsilon+\varepsilon'} \circ g \circ g'$$

puisque $(h \circ g)(\mathbf{OM}) = -g(\mathbf{OM}) = g(-\mathbf{OM}) = (g \circ h)(\mathbf{OM})$ car $g$ est linéaire. Donc

$$\varphi(f \circ f') = ([\varepsilon + \varepsilon'], \sigma_{gog'}) = ([\varepsilon] + [\varepsilon'], \sigma_g \sigma_{g'}) = \varphi(f)\varphi(f').$$

Ainsi $f$ est un homorphisme. Montrons que $\varphi$ est une bijection. Soit $([\varepsilon], \alpha) \in \mathbb{Z}_2 \times \mathcal{S}_4$ quelconque. On sait qu'il existe $g \in \mathrm{Rot}\,(C)$ unique tel que $\sigma_g = \alpha$. Alors $f = h^\varepsilon \circ g$ est l'unique élément de $\mathrm{Isom}\,(C)$ dont l'image par $\varphi$ est $([\varepsilon], \alpha)$. Tout élément de $\mathbb{Z}_2 \times \mathcal{S}_4$ a un antécédent unique par $\varphi$ donc $\varphi$ est une bijection. C'est ainsi un isomorphisme. □

**Remarque.** Tout élément $f$ de $\mathrm{Isom}\,(C)$ définit encore une permutation de l'ensemble des diagonales du cube. L'homomorphisme $\sigma : \mathrm{Rot}\,(C) \longrightarrow \mathcal{S}_4$ peut ainsi être prolongé en un homomorphisme

$$\varrho : \mathrm{Isom}\,(C) \longrightarrow \mathcal{S}_4. \tag{5.5}$$

Mais $\varrho$ *n'est pas injectif*, car le noyau de $\varrho$ est $\{\pm Id\}$.

### 5.3.2 Groupe des isométries du tétraèdre.

**Tétraèdre et cube.**

Le tétraèdre régulier $T$ de côté $\sqrt{2}$ peut être inscrit dans le cube $C$ de côté 1 en sorte que les arêtes opposées de $T$ soient des diagonales de faces opposées de $C$ (Figure 5.2).

**Lemme 5. 1.** *Une isométrie laissant $T$ invariant laisse $C$ invariant.*

**Démonstration.**

Le centre $O$ de $C$ est aussi l'isobarycentre des sommets de $T$. Soit $f$ une isométrie laissant $T$ invariant. Comme $f$ est une bijection affine, $f$ laisse aussi invariant d'après la Prop.4.7 l'ensemble des points extrêmaux de $T$ : $f(Extr(T)) = Extr(T)$. L'ensemble des points extrêmaux de $T$ est l'ensemble des sommets de $T$ d'après l'exemple (iv) de la section 4.2.6. $f$ laisse donc invariant l'ensemble des sommets de $T$, et par conséquent leur barycentre $O$ puisque $f$ est affine : $f(O) = O$. (On aurait pu aussi remarquer que pour $M, N \in T$, on a $\|\mathbf{MN}\| \leqslant \sqrt{2}$, avec égalité si et seulement si $M, N$ sont deux sommets de $T$. Comme $\|\mathbf{f(M)\,f(N)}\| = \|\mathbf{MN}\|\ \forall M, N$, cela implique que si $M, N$ sont deux sommets de $T$, il en est de même de $f(M), f(N)$).

Soit $g$ la partie linéaire de $f$. On a $f(M) = O + g(\mathbf{OM})\ \forall M$. Soit $A$ un sommet de $T$, $A'$ le sommet du cube *opposé* à $A$, *i.e.* tel que $\mathbf{OA'} = -\mathbf{OA}$. $B = f(A)$ est un sommet de $T$. Soit $B' = f(A')$. On a

$$\mathbf{OB'} = g(\mathbf{OA'}) = g(-\mathbf{OA}) = -\mathbf{OB}$$

donc $B'$ est le sommet de $C$ opposé au sommet $B$ de $T$. Il en résulte que $f$ laisse invariant l'ensemble des sommets de $C$, donc $C$ lui-même, puisque $C$ est l'enveloppe convexe de ses sommets. □

**Groupe des rotations du tétraèdre.**

Si une rotation de $E$ laisse invariant $T$, elle laisse aussi invariant $C$. Les éléments de $\mathrm{Rot}\,(T)$ sont donc les éléments de $\mathrm{Rot}\,(C)$ qui laissent invariant $T$. L'examen direct montre que ce sont, outre $Id$ :

- Les rotations d'angle $\pi$ correspondant aux axes quadruples du cube, *i.e.* aux axes joignant les milieux de deux arêtes opposées du tétraèdre.
- Les rotations d'angle $\frac{\pi}{3}$ et $\frac{2\pi}{3}$ correspondant aux axes triples du cube, *i.e.* aux axes joignant un sommet du tétraèdre au milieu de la face opposée.

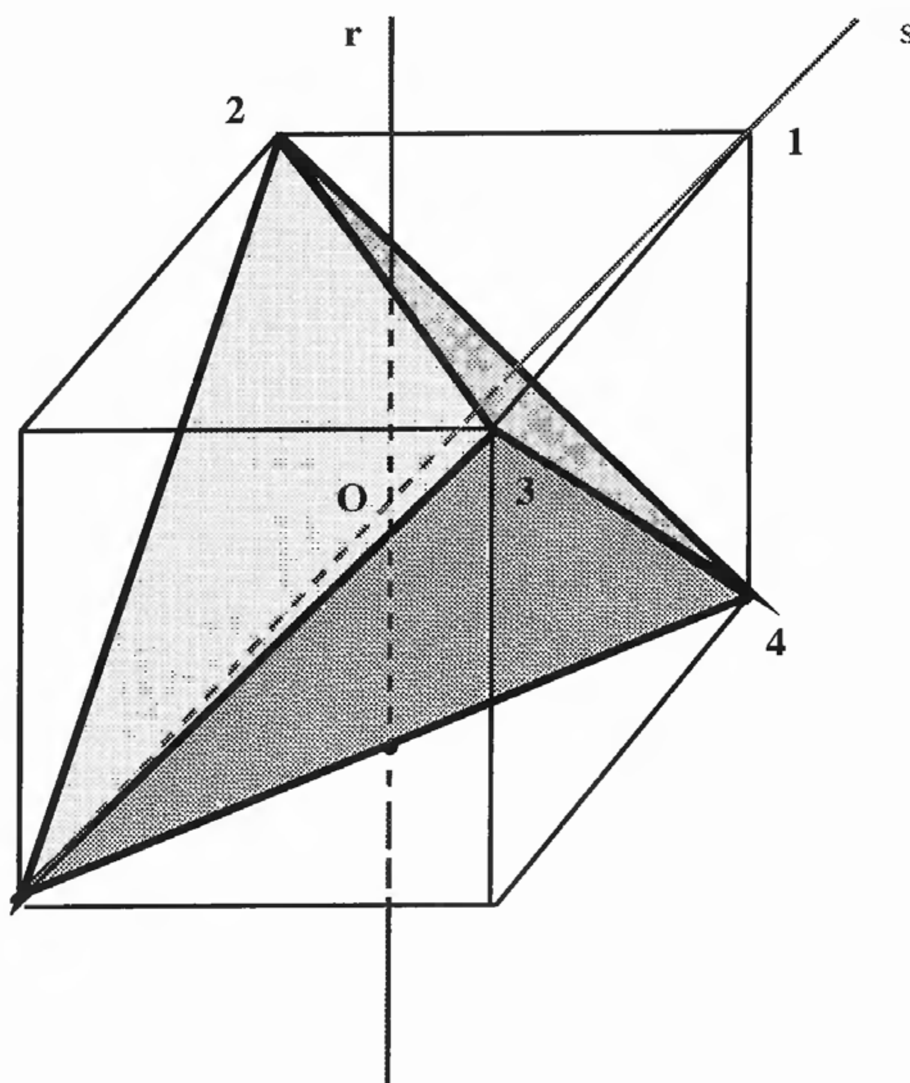



FIG. 5.2: *Les rotations du tétraèdre.*

On trouve donc pour $T$ :

| | | | |
|---|---|---|---|
| 4 | axes triples | qui donnent | $4 \cdot 2 = 8$ rotations $\neq Id$ |
| 3 | axes doubles | qui donnent | $3 \cdot 1 = 3$ rotations $\neq Id$ |
| | | | $Id = 1$ rotation |
| | | | **total = 12 rotations** |

**Théorème 5. 6.** *Soit* $\mathrm{Rot}\,(T) = \{f \in SO(\mathbb{R}^3);\ f(T) = T\}$ *l'ensemble des rotations laissant invariant le tétraèdre* $T$. $\mathrm{Rot}\,(T)$ *est un sous-groupe de* $\mathrm{Rot}\,(C)$ *isomorphe au groupe alterné* $\mathcal{A}_4$.

**Démonstration.**

Il est clair que $\mathrm{Rot}\,(T)$ est un sous-groupe de $\mathrm{Rot}\,(C)$. On a vu que $|\mathrm{Rot}\,(T)| = 12$. Considérons l'isomorphisme $\sigma : \mathrm{Rot}\,(C) \longrightarrow \mathcal{S}_4$. L'examen direct de la signature de $\sigma_f$ pour les divers éléments de $\mathrm{Rot}\,(T)$ montre que $\sigma_f \in \mathcal{A}_4$ pour $f \in \mathrm{Rot}\,(T)$. L'image de $\mathrm{Rot}\,(T)$ par l'isomorphisme $\sigma$ est donc contenue dans $\mathcal{A}_4$. Comme $\mathrm{Rot}\,(T)$ et $\mathcal{A}_4$ ont le même cardinal, la restriction de $\sigma$ à $\mathrm{Rot}\,(T)$ est un isomorphisme de $\mathrm{Rot}\,(T)$ sur $\mathcal{A}_4$. □

**Groupe des isométries du tétraèdre.**

**Théorème 5. 7.** $\mathrm{Isom}(T) = \{f \in O(\mathbb{R}^3);\ f(T) = T\}$ *est un sous-groupe de* $O(\mathbb{R}^3)$ *isomorphe au groupe symétrique* $\mathcal{S}_4$.

**Démonstration.**

Il est clair que $\mathrm{Isom}(T)$ est un sous-groupe de $\mathrm{Isom}(C)$. Identifions les sommets du tétraèdre et les diagonales du cube. Tout $f \in \mathrm{Rot}(T)$ permute les sommets. Cette permutation des sommets est, avec l'identification ci-dessus, l'élément $\sigma_f \in \mathcal{S}_4$. Maintenant, n'importe quel élément de $\mathrm{Isom}(T)$ permute aussi les sommets, donc définit un élément $\lambda_f \in \mathcal{S}_4$. On définit ainsi une application

$$\lambda : \mathrm{Isom}(T) \longrightarrow \mathcal{S}_4.$$

Cette application est un homorphisme et l'on a $\lambda_f = \sigma_f$ pour $f \in \mathrm{Rot}(T)$.

En fait, $\lambda$ n'est autre que la restriction de l'homorphisme (5.5) au sous-groupe $\mathrm{Isom}(T)$ de $\mathrm{Isom}(C)$.

L'homorphisme $\lambda$ est injectif. En effet, $\mathrm{Ker}\,\lambda = \{Id\}$ puisque si $f \in \mathrm{Isom}(T)$ laisse fixe chaque sommet, c'est forcément l'identité par linéarité.

L'homorphisme $\lambda$ est surjectif. Considérons en effet une symétrie orthogonale par rapport à un "plan médian" de $T$, *i.e.* contenant une arête et le milieu de l'arête opposée, par exemple la symétrie $J$ par rapport au plan contenant l'arête de sommets 2 et 3 et le milieu de l'arête opposée (voir Fig. 5.2). Alors $\lambda_J$ est la transposition $(1, 4)$. Soit $\nu \in \mathcal{S}_4$. Si $\nu \in \mathcal{A}_4$, on sait déjà d'après le Th. 5. 6 qu'il existe $f \in \mathrm{Rot}(T)$ tel que $\nu = \sigma_f$ donc $\nu = \lambda_f$. Si $\nu$ est *impaire*, alors $\lambda_J\,\nu$ est *paire*, *i.e.* $\lambda_J\,\nu \in \mathcal{A}_4$, donc il existe $f \in \mathrm{Rot}(T)$ tel que $\lambda_J\,\nu = \sigma_f = \lambda_f$. En composant par $\lambda_J$, on obtient

$$\nu = \lambda_J\,\lambda_f = \lambda_{J\circ f}$$

puisque $\lambda_J^2 = \lambda_{J^2} = \lambda_{Id} = Id_{\{1,2,3,4\}}$. Donc $\lambda$ est surjectif.

L'homorphisme $\lambda$ est est donc un isomorphisme. □

## 5.4 Exercices.

**Exercice 5.1.**

Montrer que $D_3$ est isomorphe à $\mathcal{S}_3$.

**Exercice 5.2.**

On considère le groupe symétrique $\mathcal{S}_n$ $(n \geqslant 2)$.

(i) Montrer que l'on a pour $i \geqslant 1$, $i + 1 < j \leqslant n$

$$(i, j) = (i, i+1)(i+1, j)(i, i+1) \tag{5.6}$$

et en déduire par récurrence que pour $1 \leqslant i < j \leqslant n$ :

$$\begin{aligned}(i, j) = (i, i+1)(i+1, i+2)\cdots(j-2, j-1)(j-1, j)\\ (j-2, j-1)\cdots(i+1, i+2)(i, i+1)\,.\end{aligned} \tag{5.7}$$

(ii) En déduire que $\{(i, i+1);\ 1 \leqslant i < n\}$ engendre $\mathcal{S}_n$.

(iii) Montrer que l'on a

$$(i, i+1) = (1, i)(1, i+1)(1, i) \quad (1 < i < n)\,. \tag{5.8}$$

(iv) En déduire en utilisant (ii) que $\{(1,i);\ 1 < i \leqslant n\}$ engendre $\mathcal{S}_n$.

(v) Soit $\tau = (1,2)$ et $\sigma = (1,2,\cdots,n)$. Montrer que

$$(i,i+1) = \sigma^{i-1}\tau\sigma^{-(i-1)} \quad (1 \leqslant i < n). \tag{5.9}$$

(vi) En déduire en utilisant (ii) que $\{\sigma,\tau\}$ engendre $\mathcal{S}_n$.

**Exercice 5.3.**

On considère un tétraèdre régulier. Combien de tétraèdres réguliers différents peut-on faire si l'on colorie chaque face d'une couleur et que l'on dispose de $k$ couleurs?

**Indication.**

Soit $G \cong \mathcal{A}_4$ le groupe des rotations laissant invariant le tétraèdre régulier, et $X$ l'ensemble des colorations du tétraèdre régulier. Le cardinal de $X$ est $|X| = k^4$. Le groupe $G$ agit sur $X$ et deux colorations $x, y \in X$ donnent *le même tétraèdre coloré* si et seulement s'il existe une rotation $R \in G$ telle que $Rx = y$. Le nombre de tétraèdres réguliers colorés différents que l'on peut faire est donc le nombre $N$ d'orbites de $X$ sous l'action de $G$ :

$$N = \frac{1}{|G|}\sum_{g\in G} X^g$$

avec $X^g = \{x \in X;\ gx = x\}$. On a successivement : $|G| = 12$; $|X^I| = |X| = k^4$; si $s$ est une rotation d'angle $\frac{2\pi}{3}$ autour d'une diagonale (axe triple), $X^s = X^{s^2} = k^2$; si $t$ est une rotation d'angle $\pi$ autour d'une droite joignant les milieux de deux arêtes opposées (axe double), $X^t = k^2$. Il y a 4 axes triples et 3 axes doubles , donc

$$N = \frac{1}{12}(4(k^2+k^2) + 3k^2 + k^4) = \frac{1}{12}(11k^2 + k^4).$$

**Exercice 5.4.**

On considère un cube. Combien de cubes différents peut-on faire si l'on colorie chaque face d'une couleur et que l'on dispose de $k$ couleurs? Cas où $k = 3$.

**Indication.**

Soit $G \cong \mathcal{S}_4$ le groupe des rotations laissant invariant le cube, et $X$ l'ensemble des colorations du cube. On a $|X| = k^6$. Le groupe $G$ agit sur $X$ et deux colorations $x, y \in X$ donnent *le même cube coloré* si et seulement s'il existe une rotation $R \in G$ telle que $Rx = y$. Le nombre de cubes réguliers colorés différents que l'on peut faire est donc le nombre $N$ d'orbites de $X$ sous l'action de $G$ :

$$N = \frac{1}{|G|}\sum_{g\in G} X^g$$

avec $X^g = \{x \in X;\ gx = x\}$. On a successivement : $|G| = 24$; $|X^I| = |X| = k^6$; si $r$ est une rotation d'angle $\frac{\pi}{2}$ autour d'une droite joignant les milieux de deux faces opposées (axe quadruple), $|X^r| = |X^{r^3}| = k^3$ et de même $|X^{r^2}| = k^4$; si $s$ est une rotation d'angle $\frac{2\pi}{3}$ autour d'une diagonale (axe triple), $|X^s| = |X^{s^2}| = k^2$; si $t$ est une rotation d'angle $\pi$ autour d'une droite joignant les milieux de deux arêtes opposées (axe double), $|X^t| = k^3$. D'où :

$$N = \frac{1}{24}(k^6 + 3k^4 + 12k^3 + 8k^2).$$

Dans le cas où $k = 3$, on trouve $N = 57$.

**Exercice 5.5.**

On considère un cube. Combien de cubes différents peut-on faire si l'on colorie chaque face, chaque arête et chaque sommet d'une couleur et que l'on dispose de $k$ couleurs?

**Indication.**

Soit $G \cong \mathcal{S}_4$ le groupe des rotations laissant invariant le cube, et $X$ l'ensemble des colorations du cube. Le cube ayant 6 faces, 12 arêtes et 8 sommets, $|X| = k^{26}$. Le groupe $G$ agit sur $X$ et deux colorations $x, y \in X$ donnent *le même cube coloré* si et seulement s'il existe une rotation $R \in G$ telle que $Rx = y$. Le nombre de cubes réguliers colorés différents que l'on peut faire est donc le nombre $N$ d'orbites de $X$ sous l'action de $G$ :

$$N = \frac{1}{|G|} \sum_{g \in G} X^g$$

avec $X^g = \{x \in X;\ gx = x\}$. On a successivement : $|G| = 24$; $|X^I| = |X| = k^{26}$; si $r$ est une rotation d'angle $\frac{\pi}{2}$ autour d'une droite joignant les milieux de deux faces opposées (axe quadruple), $|X^r| = |X^{r^3}| = (k \cdot k)^2$ (sommets et arêtes "haut" et "bas") $\cdot k$ (arêtes "verticales") $\cdot k^3$ (faces) $= k^8$ et de même $|X^{r^2}| = k^{14}$; si $s$ est une rotation d'angle $\frac{2\pi}{3}$ autour d'une diagonale (axe triple), il y a deux trièdres à considérer, et alors $|X^s| = |X^{s^2}| = (k \cdot k^2 \cdot k)^2$ (sommet fixe · arêtes et autres sommets · faces de chaque trièdre) $\cdot k^2$ (2 triplets d'arêtes restantes) $= k^{10}$; si $t$ est une rotation d'angle $\pi$ autour d'une droite joignant les milieux de deux arêtes opposées (axe double), les sommets et les faces se classent par couples, et parmi les arêtes 2 sont globalement invariantes, les 10 autres se classant en couples, donc $|X^t| = k^4$ (sommets) $\cdot k^3$ (faces) $\cdot (k^2 \cdot k^5)$ (arêtes) $= k^{14}$.

$$N = \frac{1}{24}(k^{26} + 9k^{14} + 8k^{10} + 6k^8).$$

**Exercice 5.6.**

On dispose de $k$ couleurs, et on colorie un cube en respectant la règle suivante : *si 2 faces sont opposées, on utilise la même couleur pour ces 2 faces.*

Combien de cubes différents peut-on faire?

**Indication.**

$$N = \frac{1}{24}(4k^3 + 12k^2 + 8k).$$

**Exercice 5.7.**

Combien de colliers différents à 8 perles peut on faire si l'on dispose de perles de $k$ couleurs?

**Indication.**

Un collier à 8 perles, si l'on ne tient pas compte de la couleur des perles, peut être représenté par un octogone convexe régulier $P_8$. Les perles sont les sommets. Le groupe des isométries de $\mathbb{R}^2$ laissant invariant le collier est donc le groupe diédral $D_8$ (on prend le groupe complet car on peut "retourner" le collier). Soit $X$ l'ensemble des colorations du collier, *i.e.* des agencements de perles de couleurs sur les sommets de $P_8$. On a $|X| = k^8$. Le groupe $D_8$ agit sur $X$ et deux colorations $x, y \in X$ donnent *le même collier à perles de couleurs* si et seulement si elles sont sur une même orbite

de $D_8$. Le nombre de colliers à perles de couleurs différents que l'on peut faire est donc le nombre $N$ d'orbites de $X$ sous l'action de $D_8$ :

$$N = \frac{1}{|D_8|} \sum_{g \in D_8} X^g$$

avec $X^g = \{x \in X;\ gx = x\}$. Les éléments de $D_8$ sont

$$\{I, r, r^2, r^3, r^4, r^5, r^6, r^7, s, sr, sr^2, sr^3, sr^4, sr^5, sr^6, sr^7\}$$

où $r$ est la rotation d'angle $\frac{2\pi}{8}$ et $s$ la symétrie par rapport à l'axe des $x$ (supposé passer par 2 sommets de $P_8$). On a successivement : $|D_8| = 16$; $|X^I| = |X| = k^8$; $|X^r| = |X^{r^3}| = |X^{r^5}| = |X^{r^7}| = k$; $|X^{r^2}| = |X^{r^6}| = k^2$; $|X^{r^4}| = k^4$; $|X^s| = |X^{sr^2}| = |X^{sr^4}| = |X^{sr^6}| = k^5$ (symétries par rapport aux divers axes joignant 2 sommets opposés de $P_8$); $|X^{sr}| = |X^{sr^3}| = |X^{sr^5}| = |X^{sr^7}| = k^4$ (symétries par rapport aux divers axes joignant les milieux de 2 côtés opposés de $P_8$). D'où

$$N = \frac{1}{16}(k^8 + 4k^5 + 5k^4 + 2k^2 + 4k).$$

**Exercice 5.8.**

Combien de colliers différents à 12 perles comportant 6 perles rouges (R), 2 bleues (B), 2 vertes (V) et 2 marrons (M) peut-on faire?

**Indication.**

Un collier à 12 perles, si l'on ne tient pas compte de la couleur des perles, peut être représenté par un dodécagone convexe régulier $P_{12}$. Les perles sont les sommets. Le groupe des isométries de $\mathbb{R}^2$ laissant invariant le collier est donc le groupe diédral $D_{12}$ (on prend le groupe complet car on peut "retourner" le collier). Soit $X$ l'ensemble des colorations du collier, *i.e.* des agencements de perles de couleurs sur les sommets de $P_{12}$. On a $|X| = C_{12}^6$ (choix des places des perles R) $\cdot$ $C_6^2$ (choix des places restantes pour les perles B) $\cdot$ $C_4^2$ (choix des places restantes pour les perles V) (les places restantes pour les 2 perles M sont alors bloquées) $= 24 \cdot 55 \cdot 63$. Le groupe $D_{12}$ agit sur $X$ et deux colorations $x, y \in X$ donnent *le même collier à perles de couleurs* si et seulement si elles sont sur une même orbite de $D_{12}$. Le nombre de colliers à perles de couleurs différents que l'on peut faire est donc le nombre $N$ d'orbites de $X$ sous l'action de $D_{12}$ :

$$N = \frac{1}{|D_{12}|} \sum_{g \in D_{12}} X^g$$

avec $X^g = \{x \in X;\ gx = x\}$. Les éléments de $D_{12}$ sont

$$\{I, r, r^2, \cdots, r^{11}, s, sr, \cdots, sr^{11}\}$$

où $r$ est la rotation d'angle $\frac{2\pi}{12}$ et $s$ la symétrie par rapport à l'axe des $x$ (supposé passer par 2 sommets de $P_{12}$. On a $|D_{12}| = 24$ et $|X^I| = |X|$. Soit $g = r, r^5, r^7$ ou $r^{11}$. $g$ étant d'ordre 12 (l'ordre de $r^k$ est $\frac{12}{12 \wedge k}$, voir ex. 1.13), un élément $x \in X^g$ serait une coloration avec toutes les perles d'une même couleur, ce qui est exclu. Donc $X^g = \varnothing$. Soit $g = r^2$ ou $r^{10}$. $g$ étant d'ordre 6, un élément $x \in X^g$ serait une coloration avec au moins 6 perles de chaque couleur, ce qui est impossible. Donc $X^g = \varnothing$. Soit $g = r^3$ ou $r^9$. $g$ étant d'ordre 4, un élément $x \in X^g$ serait une coloration avec au moins 4 perles de chaque couleur, ce qui est impossible. Donc $X^g = \varnothing$. Soit $g = r^4$ ou $r^8$. $g$ étant d'ordre 3, un élément $x \in X^g$ serait une coloration avec au moins 3 perles de chaque

couleur, ce qui est impossible. Donc $X^g = \varnothing$. Ainsi, le fait que les perles de certaines couleurs soient au nombre de 2 implique que $|X^g| = 0$ si $g \in D_{12}$ n'est pas d'ordre 2, *i.e.* $g \neq r^6$ ou $g \neq sr^j$ $(0 \leqslant j \leqslant 11)$. D'autre part, les perles de chaque couleur devant être en nombre pair, on voit que si $g$ est d'ordre 2, $x \in X^g$ si et seulement si les perles d'une même couleur se classent par couples de perles qui se correspondent par $g$. Il suffit de spécifier la place de la moitié des perles de chaque couleur sur 6 sommets de $P_{12}$ (dépendant de $g$) : $|X^g| = C_6^3$ (choix des places des perles R) $\cdot C_3^1$ (choix des places restantes pour les perles B) $\cdot C_2^1$ (choix des places restantes pour les perles V) (la place restante pour la perle M est alors bloquée). D'où

$$N = \frac{1}{24}(24 \cdot 55 \cdot 63 + 13 \cdot 24 \cdot 5) = 3530.$$

**Exercice 5.9.**

(i) Soit $z_k = e^{\frac{2ik\pi}{5}}$, $1 \leqslant k \leqslant 4$.

(a) Montrer que $1 + z_k + \frac{1}{z_k} + z_k^2 + \frac{1}{z_k^2} = 0$.

(b) En déduire que $2\cos\frac{2k\pi}{5}$ est une solution de l'équation $X^2 + X - 1 = 0$.

(c) Calculer $\cos\frac{2\pi}{5}$ et $\cos\frac{4\pi}{5}$.

(ii) Dans le plan affine euclidien rapporté à un repère orthonormé $(O, (\mathbf{e_1}, \mathbf{e_2}))$, soit $\mathcal{C}$ le cercle de centre $O$ et de rayon 1, $I$ le point de coordonnées $(0, 1)$, $J$ le point de coordonnées $(-\frac{1}{2}, 0)$, $\Gamma$ le cercle de centre $J$ passant par $I$, $H$ et $K$ les points d'intersection de $\Gamma$ avec l'axe des abscisses (on prend $K$ tel que son abscisse soit $> 0$).

(a) Calculer les abscisses de $H$ et $K$.

(b) En déduire une construction à la règle et au compas du pentagone convexe régulier $P_5$.

(c) Que dire du cercle $\Gamma'$ de diamètre $[I, J]$ ? En déduire une variante de cette construction.

**Indication.**

(voir Fig. 5.3). (i) (a) $1 + z_k + z_k^2 + z_k^3 + z_k^4 = \frac{1-z_k^5}{1-z_k} = 0$. Or $z_k^3 = \frac{1}{z_k^2}$ et $z_k^4 = \frac{1}{z_k}$, d'où le résultat.

(b) Posons $X_k = z_k + \frac{1}{z_k}$. Comme $\frac{1}{z_k} = \overline{z_k}$, on a $X_k = 2\,\mathfrak{Re}\; z_k = 2\cos\frac{2k\pi}{5}$. Or

$$\begin{aligned} 1 + z_k + \frac{1}{z_k} + z_k^2 + \frac{1}{z_k^2} &= 1 + z_k + \frac{1}{z_k} + \left(z_k + \frac{1}{z_k}\right)^2 - 2 \\ &= X_k^2 + X_k - 1 \end{aligned}$$

donc $X_k^2 + X_k - 1 = 0$.

(c) Les racines de $X^2 + X - 1 = 0$ sont $X = \frac{-1\pm\sqrt{5}}{2}$ donc $\cos\frac{2\pi}{5} = \frac{-1+\sqrt{5}}{4}$ et $\cos\frac{4\pi}{5} = \frac{-1-\sqrt{5}}{4}$.

(ii)(a) Le triangle $OIJ$ étant rectangle en $O$, $\|\mathbf{IJ}\| = \frac{\sqrt{5}}{2}$. Il est alors immédiat que $x_K = \frac{-1+\sqrt{5}}{2}$ et $x_H = \frac{-1-\sqrt{5}}{2}$.

(b) Soient $H_1$ et $K_1$ les milieux des segments $[O, H]$ et $[O, K]$ respectivement. On a $x_{K_1} = \frac{1}{2}x_K = \frac{-1+\sqrt{5}}{4} = \cos\frac{2\pi}{5}$ et $x_{H_1} = \frac{1}{2}x_H = \frac{-1-\sqrt{5}}{4} = \cos\frac{4\pi}{5}$. Si $\Delta_{H_1}$ et $\Delta_{K_1}$ sont les perpendiculaires à l'axe des $x$ menées par $H_1, K_1$ respectivement, les 2 points d'intersection de $\Delta_{K_1}$ avec $\mathcal{C}$ sont les sommets de $P_5$ définis par $z_1$ et $z_4 = \overline{z_1}$ et les 2 points d'intersection de $\Delta_{H_1}$ avec $\mathcal{C}$ sont les sommets de $P_5$ définis par $z_2$ et $z_3 = \overline{z_2}$. D'où la construction de $P_5$.

(c) $\Gamma'$ passe par $O$ puisque le triangle $OIJ$ est rectangle en $O$. Le centre $L$ de

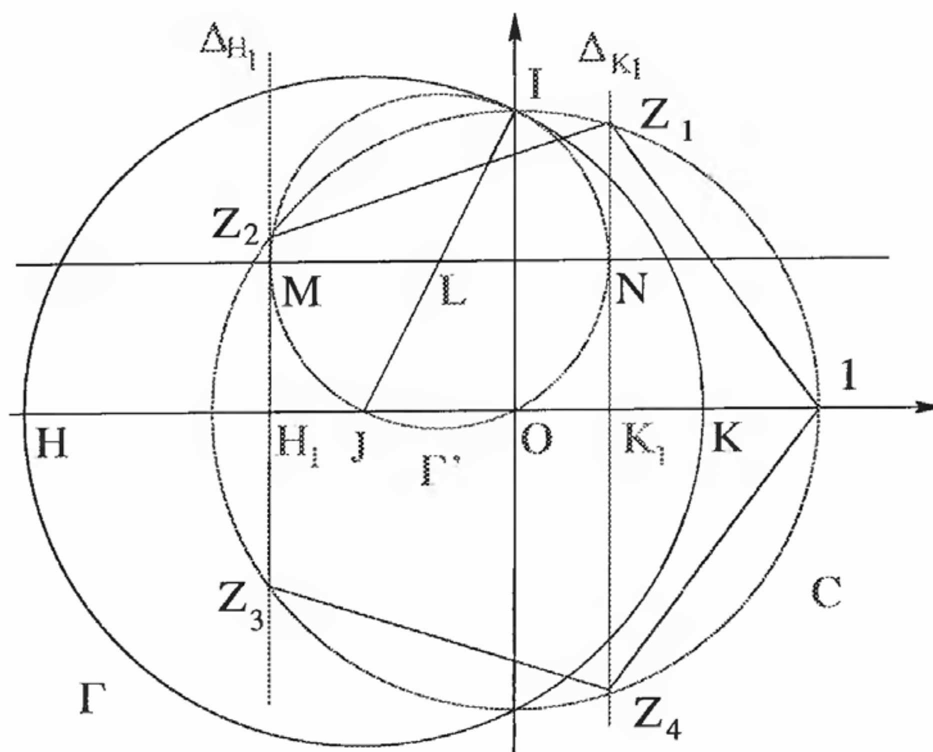



FIG. 5.3: *Construction du pentagone.*

$\Gamma'$ a pour coordonnées $(-\frac{1}{4}, \frac{1}{2})$. Son rayon est $\frac{1}{2}\|\mathbf{IJ}\| = \frac{\sqrt{5}}{4}$. La parallèle à l'axe des abscisses issue de $L$ rencontre $\Gamma'$ en deux points $M, N$ dont les abscisses sont $x_M = x_L - \frac{\sqrt{5}}{4} = \frac{-1-\sqrt{5}}{4}$, $x_N = x_L + \frac{\sqrt{5}}{4} = \frac{-1+\sqrt{5}}{4}$. La droite $\Delta_{H_1}$ passe donc par $M$ et la droite $\Delta_{K_1}$ par $N$. Or ces 2 droites sont perpendiculaires à la droite $MN$, donc tangentes au cercle $\Gamma'$ en $M$ et $N$ respectivement. Pour construire les sommets de $P_5$ définis par $z_2, z_3, z_4, z_5$, on peut donc aussi tracer le cercle $\Gamma'$, construire les 2 points $M, N$ d'intersection de $\Gamma'$ avec la parallèle à l'axe des abscisses issue de $L$ et enfin les deux perpendiculaires à l'axe des abscisses issues de $M$ et $N$. Leurs points d'intersection avec $\mathcal{C}$ sont les sommets cherchés.

**Note.** Le polygone convexe régulier à $n$ côtés $P_n$ est constructible à la règle et au compas si et seulement si $n = 2^k p_1 \cdots p_s$ où $p_1, \cdots, p_s$ sont des nombres premiers distincts et de la forme $p_i = 2^{2^{r_i}} + 1$ ($r_i \in \mathbb{N}$) (voir [6] p. 299 ou [19] p. 169 Th.17.11 ou [4] p. 51). Les entiers $n \leqslant 100$ pour lesquels $P_n$ est constructible sont

$$1, 2, 3, 4, 5, 6, 8, 10, 12, 15, 16, 17, 20, 24, 30, 32, 34, 40, 48, 51, 60, 64, 68, 80, 96.$$

**Exercice 5.10.**

Soit $G$ un groupe (noté multiplicativement) d'ordre $2n$ avec $n$ un nombre premier $\geqslant 3$.

(i) Quels sont les ordres possibles pour un élément de $G$?

(ii) Dans cette question, on suppose que tout élément de $G$ différent de l'élément neutre $e$ est d'ordre 2.

(a) Montrer que $G$ est commutatif.

(b) En considérant le sous-groupe engendré par deux éléments $a, b \in G$ différents de $e$, en déduire une contradiction.

(iii) En déduire que si $G$ n'est pas isomorphe à $\mathbb{Z}/2n\mathbb{Z}$, il existe dans $G$ un élément $a$ d'ordre $n$.

*Dans toute la suite, on suppose que $G$ n'est pas isomorphe à $\mathbb{Z}/2n\mathbb{Z}$.*

(iv) Soit $b$ un élément de $G$ n'appartenant pas au sous-groupe $H = \langle a \rangle$ engendré par $a$.
(a) Montrer que $G = H \cup bH$.
(b) En déduire que $b^2 \in H$.
(c) En raisonnant par l'absurde, montrer que $b^2 = e$.

(v) Que dire de $G$?

**Indication.**

(i) Les ordres possibles sont $1, 2, n$ et $2n$.
(ii)(a) Pour tout $x \in G$, on a $x^2 = e$ donc $x^{-1} = x$. Alors pour tous $x, y \in G$, $xy = (xy)^{-1} = y^{-1}x^{-1} = yx$.
(b) Comme $G$ est commutatif tel que $x^2 = e\ \forall x \in G$, $\{e, a, b, ab\}$ est un sous-groupe, donc c'est $\langle a, b \rangle$. Or il est d'ordre 4, ce qui est impossible.
(iii) Il existe donc dans $G$ un élément $a \neq e$ dont l'ordre n'est pas 2. D'après (i), l'ordre de $a$ est $n$ ou $2n$. Mais si c'était $2n$, $G$ serait engendré par $a$ donc isomorphe à $\mathbb{Z}/2n\mathbb{Z}$. Ainsi $a$ est d'ordre $n$.
(iv) (a) Il y a 2 classes modulo $H$ puisque $|G/H| = |G|/|H| = 2$. Or $b \notin H$ donc $bH \neq H$. Les deux classes sont donc $H$ et $bH$, et ainsi $G = H \cup bH$.
(b) La condition $b^2 \notin H$ impliquerait $b^2 \in bH$ d'après (a), donc en multipliant par $b^{-1}$, $b \in H$, ce qui est contradictoire. Ainsi $b^2 \in H$.
(c) Supposons $b^2 \neq e$. L'ordre de $b$ est donc $n$ ou $2n$. Or $G$ n'étant pas isomorphe à $\mathbb{Z}/2n\mathbb{Z}$, il n'y a pas d'élément d'ordre $2n$, donc l'ordre de $b$ est $n$. Ainsi $b^n = e$. Or $n$ est impair. Soit $q$ tel que $n = 2q + 1$. Comme $b^2 \in H = \langle a \rangle$, il existe $i$, $0 < i < n$ tel que $b^2 = a^i$. Alors

$$b^n = e \Leftrightarrow b\,(b^2)^q = e \Leftrightarrow b\,(a^i)^q = e \Leftrightarrow ba^{iq} = e \Leftrightarrow b = a^{-iq}$$

donc $b \in H$. Comme par hypothèse $b \notin H$, il y a contradiction. Cela prouve que $b^2 = e$.
(v) D'après (iv) tout élément de $G$ n'appartenant pas à $H$ est d'ordre 2. Soit $b$ un tel élément. D'après (iv)(a), $G = \langle a, b \rangle$. De plus, $ab \notin H$, donc $(ab)^2 = e$ *i.e.* $abab = e$ ou encore $bab = a^{-1}$. On en déduit que $G$ est isomorphe à $D_n$, d'après le Th.5.3.

**Exercice 5.11.**

**Solides platoniciens.**

$\Sigma$ désigne un polyèdre convexe de l'espace à 3 dimensions. On suppose $\Sigma$ régulier, *i.e.* vérifiant simultanément les 2 conditions suivantes.

- Les faces de $\Sigma$ sont des polygones réguliers à $p$ côtés ($p \geq 3$);
- En chaque sommet de $\Sigma$ converge le même nombre $q$ d'arêtes ($q \geq 3$).

Soient $s$ le nombre de sommets, $a$ le nombre d'arêtes et $f$ le nombre de faces de $\Sigma$.

On admet sans démonstration la formule d'Euler : $s - a + f = 2$.
(i) Montrer que $pf = 2a$.
(ii) Montrer que $qs = 2a$.
(iii) En utilisant la formule d'Euler, montrer que

$$(s/2p)[4 - (p-2)(q-2)] = 2$$

et en déduire que $(p-2)(q-2) < 4$

(iv) Compléter le tableau suivant en supprimant toute ligne inutile.

| p | q | $(p-2)(q-2)$ | s | a | f | Nom du polyèdre $\Sigma$ |
|---|---|---|---|---|---|---|
| | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |

Combien de sortes de polyèdres trouve-t-on? Parmi les polyèdres trouvés, certains couples sont-ils "duaux" en un sens à préciser, et lesquels?

**Indication.**

(i) Chaque arête est commune à deux faces, donc le nombre de côtés d'une face multiplié par le nombre de faces donne le *double* du nombre d'arêtes.

(ii) Chaque arête est issue de deux sommets, donc le nombre de sommets multiplié par le nombre d'arêtes issues d'un sommet donne le *double* du nombre d'arêtes.

(iii)

$$pf = qs = 2a$$

$$a = qs/2$$

$$f = qs/p$$

$$s - a + f = s - (qs)/2 + (qs)/p = (s/2p)(2p - pq + 2q)$$

Or $2p-pq+2q = -p(q-2)+2q = -(p-2)(q-2)-2(q-2)+2q = -(p-2)(q-2)+4$.

Donc

$$(s/2p)[4 - (p-2)(q-2)] = 2.$$

Alors $s/2p > 0 \Rightarrow (p-2)(q-2) < 4$ et

$$s = 4p/[4 - (p-2)(q-2)].$$

$$a = qs/2$$

$$f = 2a/p$$

| p | q | $(p-2)(q-2)$ | s | a | f | Nom du polyédre $\Sigma$ |
|---|---|---|---|---|---|---|
| 3 | 3 | 1 | 4 | 6 | 4 | tétraèdre |
| 3 | 4 | 2 | 6 | 12 | 8 | octaèdre |
| 3 | 5 | 3 | 12 | 30 | 20 | icosaèdre |
| 4 | 3 | 2 | 8 | 12 | 6 | cube |
| 5 | 3 | 3 | 20 | 30 | 12 | dodécaèdre |
| //// | /// | /////// | /// | /// | /// | ////////// |

Il y a donc 5 solides platoniciens. Si l'on appelle *dual* d'un polyèdre le polyèdre obtenu en joignant les centres de ses faces, le cube et l'octaèdre forment un couple dual, le dodécaèdre et l'icosaèdre (voir Fig.5.8 et 5.9) un autre couple dual et le

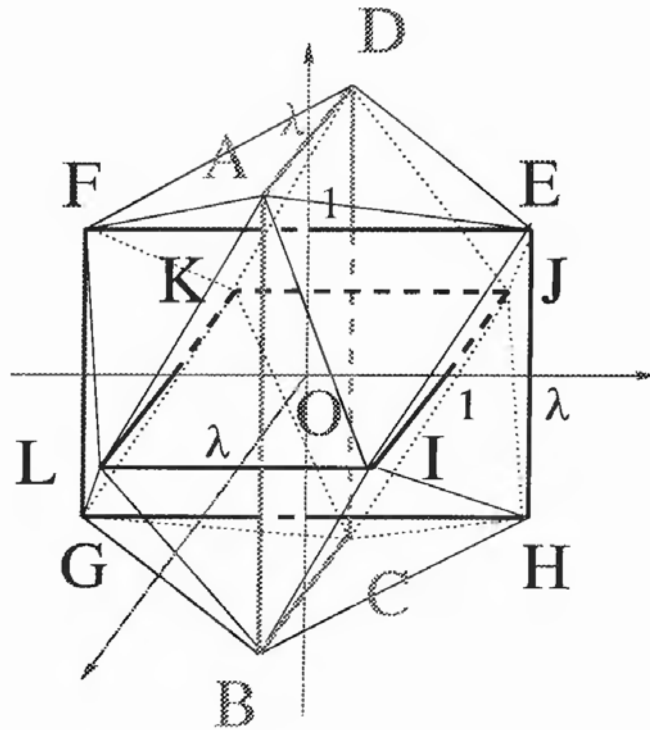



FIG. 5.4: *Trois rectangles d'or orthogonaux et l'icosaèdre.*

tétraèdre est self-dual. Pour ce qui concerne les solides platoniciens et les polyèdres en général, on pourra consulter [7] qui contient de nombreuses représentations et aspects historiques. Mentionnons simplement que Platon pensait que les 5 polyèdres réguliers étaient les constituants de l'univers physique. Il associait chaque polyèdre avec un élément : le cube avec la terre, le tétraèdre avec le feu, l'octaèdre avec l'air, l'icosaèdre avec l'eau, et le dodécaèdre avec le 5ème élément qui était pour Platon l'univers.

**Exercice 5.12.**

Dans l'espace affine euclidien de dimension 3 rapporté à un repère orthonormé $(O, (\mathbf{e_1}, \mathbf{e_1}, \mathbf{e_3}))$, soient, pour $\lambda > 0$ fixé, les 12 points $A = \begin{pmatrix} 1 \\ 0 \\ \lambda \end{pmatrix}, B = \begin{pmatrix} 1 \\ 0 \\ -\lambda \end{pmatrix}$,
$C = \begin{pmatrix} -1 \\ 0 \\ -\lambda \end{pmatrix}, D = \begin{pmatrix} -1 \\ 0 \\ \lambda \end{pmatrix}, E = \begin{pmatrix} 0 \\ \lambda \\ 1 \end{pmatrix}, F = \begin{pmatrix} 0 \\ -\lambda \\ 1 \end{pmatrix}, G = \begin{pmatrix} 0 \\ -\lambda \\ -1 \end{pmatrix}, H = \begin{pmatrix} 0 \\ \lambda \\ -1 \end{pmatrix}$,
$I = \begin{pmatrix} \lambda \\ 1 \\ 0 \end{pmatrix}, J = \begin{pmatrix} -\lambda \\ 1 \\ 0 \end{pmatrix}, K = \begin{pmatrix} -\lambda \\ -1 \\ 0 \end{pmatrix}, L = \begin{pmatrix} \lambda \\ -1 \\ 0 \end{pmatrix}$ et $Y$ le polyèdre ayant pour sommets ces 12 points (voir Fig. 5.4 et 5.5). Montrer que $Y$ est un icosaèdre régulier si et seulement si $\lambda = \frac{1+\sqrt{5}}{2}$.

**Indication.**

$Y$ est un icosaèdre régulier si et seulement si ses faces sont des triangles équilatéraux. Il suffit pour cela que

$$\|\mathbf{AD}\| = \|\mathbf{AE}\| = \|\mathbf{AI}\| = \|\mathbf{EI}\|.$$

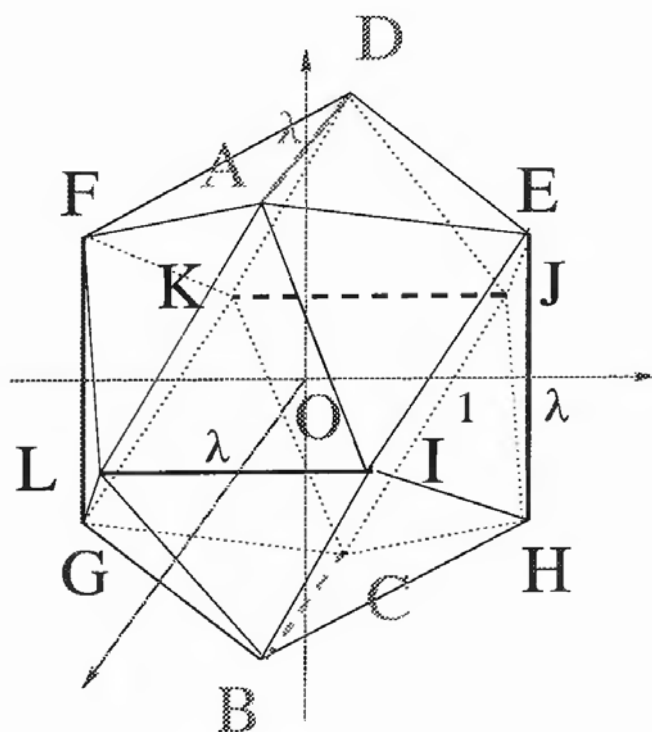



FIG. 5.5: *L'icosaèdre.*

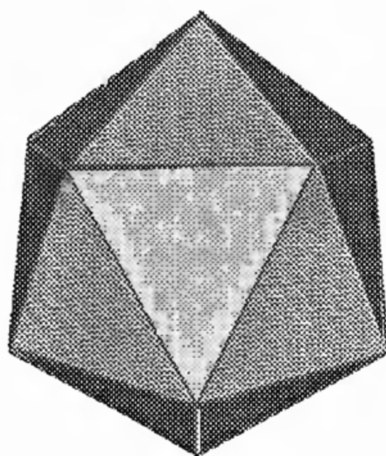


FIG. 5.6: *Une vue de l'icosaèdre.*

Or $\|\mathbf{AD}\|^2 = 4$ et

$$\|\mathbf{AE}\|^2 = \|\mathbf{AI}\|^2 = \|\mathbf{EI}\|^2 = 1 + \lambda^2 + (\lambda - 1)^2.$$

La condition est donc $\lambda^2 - \lambda - 1 = 0$. L'unique solution $> 0$ de cette équation est

$$\lambda = \frac{1+\sqrt{5}}{2}. \tag{5.10}$$

(5.10) est appelé *nombre d'or*. On a $\lambda \approx 1.6180$.

**Exercice 5.13.**

Soient $A, B, C$ 3 points d'une sphère de centre $O$ de l'espace affine euclidien de dimension 3, et $G$ leur isobarycentre. Montrer que si le triangle $ABC$ est équilatéral, $\mathbf{OG}$ est orthogonal au plan $ABC$.

**Exercice 5.14.**

(i) Montrer que toute isométrie de l'espace affine $\mathbb{R}^3$ laissant invariant l'icosaèdre $Y$ laisse aussi invariant son centre $O$.

(ii) Soit $\mathrm{Rot}\,(Y)$ l'ensemble des rotations laissant invariant $Y$. Montrer que $\mathrm{Rot}\,(Y)$ est un groupe de cardinal 60.

(iii) Montrer que $\mathrm{Rot}\,(Y)$ est isomorphe au groupe alterné $\mathcal{A}_5$.

**Indication.**

(i) L'icosaèdre $Y$ est invariant par la symétrie de centre $O$. Les 12 sommets forment 6 couples de sommets opposés. On peut donc effectuer le même raisonnement que pour le cube.

(ii) On vérifie immédiatement que $\mathrm{Rot}\,(Y)$ est un groupe. Montrons que $|\mathrm{Rot}\,(Y)| = 60$. L'icosaèdre $Y$ étant invariant par la symétrie de centre $O$, ses 20 faces forment 10 couples de faces symétriques. Pour chaque couple, la droite joignant les isobarycentres des deux faces passe par $O$ et c'est un axe d'ordre 3. De même, les 30 arêtes forment 15 couples d'arêtes opposées. La droite joignant les milieux de deux arêtes opposées est un axe double. Enfin, les 12 sommets forment 6 couples de sommets opposés, et la droite joignant deux sommets opposés est un axe d'ordre 5 (voir Fig. 5.7). On trouve donc pour $Y$ :

| | | | |
|---|---|---|---|
| 6 | axes d'ordre 5 | qui donnent | $6 \cdot 4 = 24$ rotations $\neq Id$ |
| 10 | axes triples | qui donnent | $10 \cdot 2 = 20$ rotations $\neq Id$ |
| 15 | axes doubles | qui donnent | $15 \cdot 1 = 15$ rotations $\neq Id$ |
| | | | $Id = 1$ rotation |
| | | | **total = 60 rotations** |

(iii) Chacun des axes doubles passe par les milieux de deux arêtes opposées. Ces axes doubles viennent en ensembles de 3 axes deux-à-deux orthogonaux. Les 15 axes doubles se classent donc en 5 tels ensembles de 3 axes $\{t_1, t_2, t_3, t_4, t_5\}$. A chaque $t_i$ $(1 \leqslant i \leqslant 5)$ correspond un ensemble de 3 couples d'arêtes opposées, chaque couple correspondant à l'un des axes doubles de $t_i$, et réciproquement. Soit $f \in \mathrm{Rot}\,(\mathrm{Y})$. Si $\Delta_1, \Delta_2, \Delta_3$ sont 3 axes orthogonaux deux-à-deux et passant chacun par les milieux de 2 arêtes opposées, alors $f(\Delta_1), f(\Delta_2), f(\Delta_3)$ est un ensemble de 3 axes qui a les mêmes propriétés. $f$ définit donc une application $\Phi_f$ de $\{t_1, t_2, t_3, t_4, t_5\}$ dans lui-même. Cette application est injective : pour $t = \{\Delta_1, \Delta_2, \Delta_3\}$ et $t' = \{\Delta'_1, \Delta'_2, \Delta'_3\}$ l'équation $\Phi_f(t) = \Phi_f(t')$ s'écrit $\{f(\Delta_1), f(\Delta_2), f(\Delta_3)\} = \{f(\Delta'_1), f(\Delta'_2), f(\Delta'_3)\}$, donc comme $f$ est bijective, les ensembles $t$ et $t'$ sont les mêmes. L'application $\Phi_f$

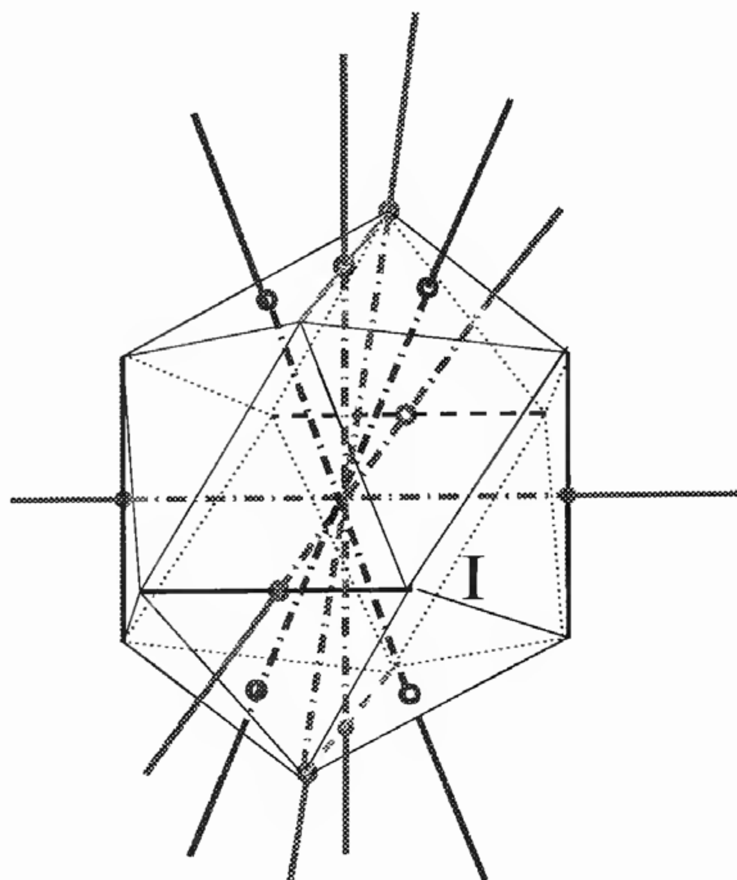



FIG. 5.7: *Quelques axes de l'icosaèdre.*

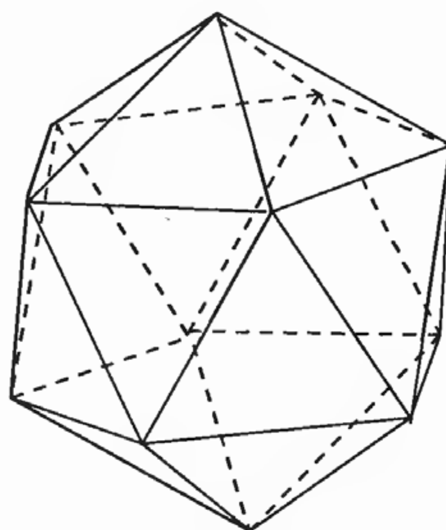


FIG. 5.8: *Encore l'icosaèdre.*

est alors une bijection de $\{t_1, t_2, t_3, t_4, t_5\}$ sur lui-même, *i.e.* un élément de $\mathcal{S}_5$. Il est enfin immédiat que l'application

$$\Phi \ : \ \mathrm{Rot}\,(\mathrm{Y}) \to \mathcal{S}_5 \tag{5.11}$$

définie par $f \mapsto \Phi_f$ est un homomorphisme de groupes. On va montrer que $\Phi$ est un isomorphisme de Rot (Y) sur $\mathcal{A}_5$. Vérifions d'abord que l'image de $\Phi$ est contenue dans $\mathcal{A}_5$. Soit $f \in \mathrm{Rot}\,(\mathrm{Y}), f \neq Id$. On va calculer $\Phi_f$. Il y a 3 cas possibles pour l'axe $\Delta$ de $f$.

• $\Delta$ est un axe d'ordre 5 de $Y$. Dans ce cas, $\Delta$ passe par 2 sommets opposés de $Y$. Soit $A$ l'un de ces 2 sommets. Il y a 5 arêtes issues de $A$. Elles correspondent à des sommets $E, D, F, L, I$ (voir Fig.5.5). Chacune de ces 5 arêtes correspond à un ensemble de 3 couples d'arêtes opposées, donc à l'un des éléments $t_1, t_2, t_3, t_4, t_5$. Deux quelconques de ces 5 arêtes ne sont ni orthogonales ni parallèles. La correspondance ci-dessus est donc une bijection entre l'ensemble de ces 5 arêtes et $\{t_1, t_2, t_3, t_4, t_5\}$. Or $f$ permute cycliquement les sommets $E, D, F, L, I$ et les arêtes correspondantes puisque son angle est de la forme $\frac{2k\pi}{5}, (1 \leqslant k < 5)$. Donc $\Phi_f$ est un 5-cycle de la forme $(t_{i_1}, t_{i_2}, t_{i_3}, t_{i_4}, t_{i_5})$.

• $\Delta$ est un axe d'ordre 3 de $Y$. Dans ce cas, $\Delta$ passe par les centres de 2 faces opposées. L'angle de rotation de $f$ est alors $\frac{2\pi}{3}$ ou $\frac{4\pi}{3}$. En choisissant une orientation de $\Delta$, on peut supposer que c'est $\frac{2\pi}{3}$. Fixons l'une des faces $ADF$ par laquelle passe $\Delta$. Deux quelconques de ses 3 arêtes ne sont ni orthogonales ni parallèles. Les 3 arêtes $AD, DF, FA$ correspondent donc respectivement à 3 éléments $t_{i_1}, t_{i_2}, t_{i_3}$ deux-à-deux distincts. De $A$ sont issues 5 arêtes, dont $AD$ et $AF$. Il reste 3 arêtes que nous noterons $AL, AI, AE$ comme sur la figure 5.5. $AI$ est orthogonale à $DF$ (En effet, $A$ et $I$ sont dans le plan médiateur de $[L, E]$. Or **LE** est parallèle à **FD** car $EDFLI$ est un pentagone régulier. Donc les plans médiateurs de $[L, E]$ et$[F, D]$ ont même direction. Comme $A$ est un point commun, ces deux plans coïncident). Il s'ensuit que $AI$ donne le même élément $t_{i_2}$ que $DF$. Les 2 arêtes $AL$ et $AE$ correspondent alors respectivement aux 2 éléments $t_{i_4}$ et $t_{i_5}$ restants. Or $f$ permute cycliquement les sommets $A, D, F$. Donc la restriction de $\Phi_f$ à $\{t_{i_1}, t_{i_2}, t_{i_3}\}$ est le 3-cycle $(t_{i_1}, t_{i_2}, t_{i_3})$. D'autre part, l'image par $f$ de $AL$ est $DE$. Mais $AL$ est orthogonal à $DE$ (pour la même raison que $AI$ et $FD$). Donc $\Phi_f(t_{i_4}) = t_{i_4}$. Alors $\Phi_f(t_{i_5}) = t_{i_5}$. Donc

$$\Phi_f = (t_{i_1}, t_{i_2}, t_{i_3})\ .$$

• $\Delta$ est un axe d'ordre 2 de $Y$. Dans ce cas cet axe fait partie de l'un de $t_i$, soit $t_{i_1}$, et l'on a $\Phi_f(t_{i_1}) = t_{i_1}$. Soit $AD$ une arête par le milieu de laquelle passe l'axe. Les deux faces contigues à $AD$ déterminent 4 autres arêtes notées comme sur la Fig. 5.5 $AF, DE, DF, AE$. Elles correspondent respectivement aux 4 éléments $t_{i_2}, t_{i_3}, t_{i_4}, t_{i_5}$ restants. Or $f(AF) = DE$ et $f(DF) = AE$ donc $\Phi_f(t_{i_2}) = t_{i_3}$ et $\Phi_f(t_{i_4}) = t_{i_5}$. Alors $\Phi_f = (t_{i_2}, t_{i_3})(t_{i_4}, t_{i_5})$.

On constate donc que dans tous le cas $\Phi_f \in \mathcal{A}_5$. De plus, on constate que pour $f \neq Id$ , on a $\Phi_f \neq Id$ , *i.e.* $\mathrm{Ker}\,\Phi = \{Id\}$. Donc $\Phi$ est injective. Comme $|\mathrm{Rot}\,(\mathrm{Y})| = |\mathcal{A}_5|$ , $\Phi$ est un isomorphisme de Rot (Y) sur $\mathcal{A}_5$ .

**Exercice 5.15.**

Soit $\mathrm{Isom}\,(Y) = \{f \in O(\mathbb{R}^3);\ f(Y) = Y\}$ l'ensemble des isométries laissant invariant l'icosaèdre $Y$. Montrer que Isom $(Y)$ est un sous-groupe de $O(\mathbb{R}^3)$ isomorphe au groupe produit direct

$$\mathbb{Z}_2 \times \mathcal{A}_5.$$

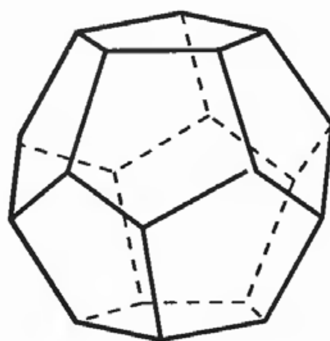


FIG. 5.9: *Le dodécaèdre.*

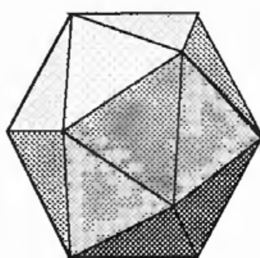


FIG. 5.10: *Logo A.*

**Indication.**

Raisonner comme pour le cube en utilisant la symétrie $h$ par rapport à l'origine et l'ex. 5.14.

**Exercice 5.16.**

Lequel des 2 logos $A$ ou $B$ des Fig. 5.10 et 5.11 est-il mathématiquement impossible pour une représentation de l'icosaèdre? (Il s'agit du logo de la *MAA, Mathematical Association of America.* Cet exercice s'inspire de [5]).

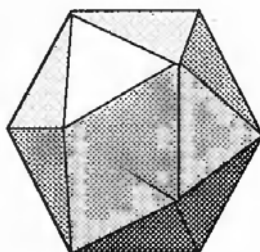


FIG. 5.11: *Logo B.*

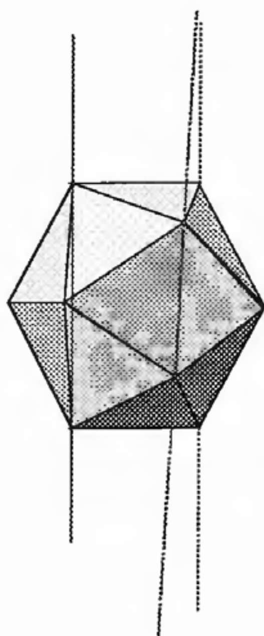


FIG. 5.12:

**Indication.**

Si le logo $A$ était mathématiquement correct, dans la Fig. 5.12, les 3 droites seraient parallèles dans l'espace. Si elles sont représentées en projection, elles restent parallèles. Si elles sont représentées en perspective, elles sont concourantes au *point de fuite*. Le logo $A$ est donc impossible mathématiquement. Le logo correct est le logo $B$ (Fig. 5.13).

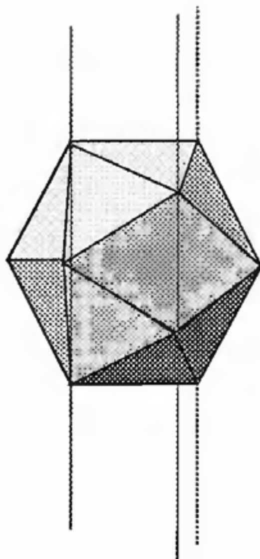


FIG. 5.13:

# Chapitre 6

# Géométrie euclidienne plane.

## 6.1 Coniques.

### 6.1.1 Ellipse.

**Définition bifocale.**

Soit $E$ le plan affine euclidien, $F, F'$ deux points distincts de $E$, $O$ le milieu du segment $[F, F']$, et $c = \|\mathbf{OF}\| > 0$.

Soit $\mathcal{E} = \{M \in E;\ \|\mathbf{MF}\| + \|\mathbf{MF'}\| = 2a\}$. Si $\mathcal{E} \neq \varnothing$ et $M \in \mathcal{E}$, on a d'après l'inégalité triangulaire $\|\mathbf{FF'}\| \leqslant \|\mathbf{MF}\| + \|\mathbf{MF'}\|$, *i.e.* $c \leqslant a$. Donc si $a < c$, $\mathcal{E} = \varnothing$. Si $a = c$ on a $\mathcal{E} = [F, F']$, et si $a > c$ il y a deux points de la médiatrice du segment $[F, F']$ qui appartiennent à $\mathcal{E}$.

**Définition 6. 1.** *Soit $E$ le plan affine euclidien, $F, F'$ deux points distincts de $E$, $O$ le milieu du segment $[F, F']$, $c = \|\mathbf{OF}\| > 0$, et $a > c$. On appelle ellipse de foyers $F, F'$, de grand axe $a$ et de distance focale $c$ l'ensemble*

$$\mathcal{E} = \{M \in E;\ \|\mathbf{MF}\| + \|\mathbf{MF'}\| = 2a\}.$$

On prend comme origine du plan le point $O$ et on utilise un repère orthonormé affine $(O, (\mathbf{e_1}, \mathbf{e_2}))$ avec $\mathbf{e_1} = \frac{\mathbf{OF}}{\|\mathbf{OF}\|}$, et $\mathbf{e_2}$ vecteur déduit de $\mathbf{e_1}$ par rotation d'angle $\frac{\pi}{2}$.

Alors $\mathbf{OM} = x\mathbf{e_1} + y\mathbf{e_2} = \begin{pmatrix} x \\ y \end{pmatrix}$. Nous identifions le point $M$ au vecteur $\mathbf{OM}$ et noterons aussi $M = \begin{pmatrix} x \\ y \end{pmatrix}$.

L'axe focal est la droite $FF'$, de vecteur directeur $\mathbf{e_1}$ (*axe des* $x$), l'axe non focal est la droite perpendiculaire en $O$, de vecteur directeur $\mathbf{e_2}$ (*axe des* $y$). $O$ est centre de symétrie de $\mathcal{E}$. En effet, si $s$ désigne la symétrie par rapport à $O$, $F' = s(F)$ et si l'on pose $M' = s(M)$ on a $\|\mathbf{M'F'}\| = \|\mathbf{MF}\|$ et $\|\mathbf{M'F}\| = \|\mathbf{MF'}\|$ donc $\|\mathbf{M'F}\| + \|\mathbf{M'F'}\| = \|\mathbf{MF}\| + \|\mathbf{MF'}\|$ et ainsi $M' \in \mathcal{E} \Leftrightarrow M \in \mathcal{E}$. Les axes focal et non focal sont des axes de symétrie de $\mathcal{E}$. Il y a deux points $B, B'$ symétriques par rapport à $O$ sur l'axe non focal appartenant à $\mathcal{E}$ : leur distance commune à $F$ ou $F'$ est $a$. Si $b$ désigne la distance de $B$ ou $B'$ à $O$, on a $a^2 = b^2 + c^2$ puisque le triangle $OFB$ est rectangle en $O$. $b$ s'appelle le *petit axe*. Il est tel que $0 < b < a$. Il y a deux points $A, A'$ symétriques par rapport à $O$ sur l'axe focal appartenant à $\mathcal{E}$ : ce sont les deux points d'abscisse $\pm a$.

**Équation cartésienne.**

**Théorème 6. 1.** *L'équation cartésienne de l'ellipse $\mathcal{E}$ de grand axe $a$ et petit axe $b$ est dans le repère orthonormé $(O,(\mathbf{e_1},\mathbf{e_2}))$ :*

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1. \tag{6.1}$$

**Démonstration.**

Dans la base orthonormée choisie, on a $\mathbf{OF} = \begin{pmatrix} c \\ 0 \end{pmatrix}$ et $\mathbf{OM} = \begin{pmatrix} x \\ y \end{pmatrix}$ donc $\mathbf{FM} = \begin{pmatrix} x-c \\ y \end{pmatrix}$ et $\|\mathbf{FM}\| = \sqrt{(x-c)^2+y^2}$. De même $\mathbf{F'M} = \begin{pmatrix} x+c \\ y \end{pmatrix}$ et $\|\mathbf{F'M}\| = \sqrt{(x+c)^2+y^2}$. Ainsi :

$$\begin{aligned}
M\in\mathcal{E} &\Leftrightarrow \|\mathbf{FM}\|+\|\mathbf{F'M}\|=2a\\
&\Leftrightarrow \|\mathbf{FM}\|^2+\|\mathbf{F'M}\|^2+2\|\mathbf{FM}\|\|\mathbf{F'M}\|=4a^2\\
&\Leftrightarrow 2(x^2+y^2+c^2)+2\|\mathbf{FM}\|\|\mathbf{F'M}\|=4a^2\\
&\Leftrightarrow \|\mathbf{FM}\|\|\mathbf{F'M}\|=2a^2-(x^2+y^2+c^2)\\
&\Leftrightarrow \begin{cases}\|\mathbf{FM}\|^2\|\mathbf{F'M}\|^2=4a^4-4a^2(x^2+y^2+c^2)+(x^2+y^2+c^2)^2\\ x^2+y^2+c^2\leqslant 2a^2.\end{cases}
\end{aligned}$$

Mais

$$\begin{aligned}
\|\mathbf{FM}\|^2\|\mathbf{F'M}\|^2 &= (x^2+y^2+c^2-2xc)(x^2+y^2+c^2+2xc)\\
&= (x^2+y^2+c^2)^2-4x^2c^2,
\end{aligned}$$

donc

$$\begin{aligned}
M\in\mathcal{E} &\Leftrightarrow \begin{cases}-4x^2c^2=4a^4-4a^2(x^2+y^2+c^2)\\ x^2+y^2+c^2\leqslant 2a^2\end{cases}\\
&\Leftrightarrow \begin{cases}x^2(a^2-c^2)+a^2y^2=a^2(a^2-c^2)\\ x^2+y^2+c^2\leqslant 2a^2\end{cases}\\
&\Leftrightarrow \begin{cases}\frac{x^2}{a^2}+\frac{y^2}{b^2}=1\\ x^2+y^2+c^2\leqslant 2a^2\end{cases}\\
&\Leftrightarrow \frac{x^2}{a^2}+\frac{y^2}{b^2}=1
\end{aligned}$$

puisque la condition $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ implique $x^2+\frac{a^2}{b^2}y^2=a^2$ donc $x^2+y^2\leqslant a^2$ (car $b<a$) et $x^2+y^2+c^2\leqslant a^2+c^2<2a^2$. □

**Paramétrage.**

**Définition 6. 2.** *Le cercle $C$ de centre $O$ et de rayon $a$ est appelé le cercle principal de l'ellipse $\mathcal{E}$.*

L'équation $x^2+y^2=a^2$ du cercle principal s'écrivant aussi

$$\frac{x^2}{a^2}+\frac{y^2}{a^2}=1\,,$$

on voit que l'ellipse se déduit du cercle principal par l'application

$$M = \begin{pmatrix} x \\ y \end{pmatrix} \mapsto M' = \begin{pmatrix} x' \\ y' \end{pmatrix} \tag{6.2}$$

définie par

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{b}{a} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

(*affinité orthogonale d'axe Ox et de rapport* $\frac{b}{a}$).

Or l'application

$$t \mapsto M(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

$$x(t) = a\cos t, \quad y(t) = a\sin t \tag{6.3}$$

de $[0, 2\pi[$ sur le cercle principal $\mathcal{C}$ est une bijection qui définit donc un paramétrage de $\mathcal{C}$. Par composition avec l'application (6.2), on donc obtient la bijection suivante de $[0, 2\pi[$ sur $\mathcal{E}$ (voir Fig. 6.1) :

$$t \mapsto M(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

$$x(t) = a\cos t, \quad y(t) = b\sin t \tag{6.4}$$

Dans ce paramétrage de l'ellipse, le paramètre $t$ s'appelle *l'angle excentrique.*

**Remarque.** Un cercle $x^2+y^2 = a^2$ peut être considéré comme une ellipse dégénérée dont les deux foyers sont confondus en $O$, avec $a = b$ et $c = 0$.

**Définition par foyer-directrice.**

**Définition 6. 3.** *Soit $\mathcal{E}$ une ellipse.*
*(i) On appelle excentricité de $\mathcal{E}$ le rapport*

$$e = \frac{c}{a} \qquad 0 < e < 1.$$

*(ii) Soit $F$ un foyer de $\mathcal{E}$. On appelle directrice relative au foyer $F$ la perpendiculaire à l'axe focal au point d'abscisse $x = \frac{a^2}{c} = \frac{a}{e}$.*

**Théorème 6. 2.** *Soit $\mathcal{E}$ une ellipse, $e$ son excentricité, $F$ un foyer et $\mathcal{D}$ la directrice relative au foyer $F$.*

*Alors $\mathcal{E} = \{M \in E;\ \frac{\|\mathbf{MF}\|}{\|\mathbf{MH}\|} = e\}$, où le point $H$ est le projeté orthogonal de $M$ sur $\mathcal{D}$, i.e. $\|\mathbf{MH}\|$ est la distance du point $M$ à $\mathcal{D}$.*

**Démonstration.**

On a $\|\mathbf{MF}\|^2 = (x-c)^2 + y^2$ et $\|\mathbf{MH}\|^2 = (x - \frac{a^2}{c})^2$.

L'équation $\|\mathbf{MF}\|^2 = e^2\|\mathbf{MH}\|^2$ s'écrit donc

$$(x-c)^2 + y^2 = \frac{c^2}{a^2}(x - \frac{a^2}{c})^2,$$

soit en développant :

$$x^2(a^2-c^2) + y^2a^2 = a^2(a^2-c^2).$$

C'est l'équation (6.1) de l'ellipse, d'où le résultat. □

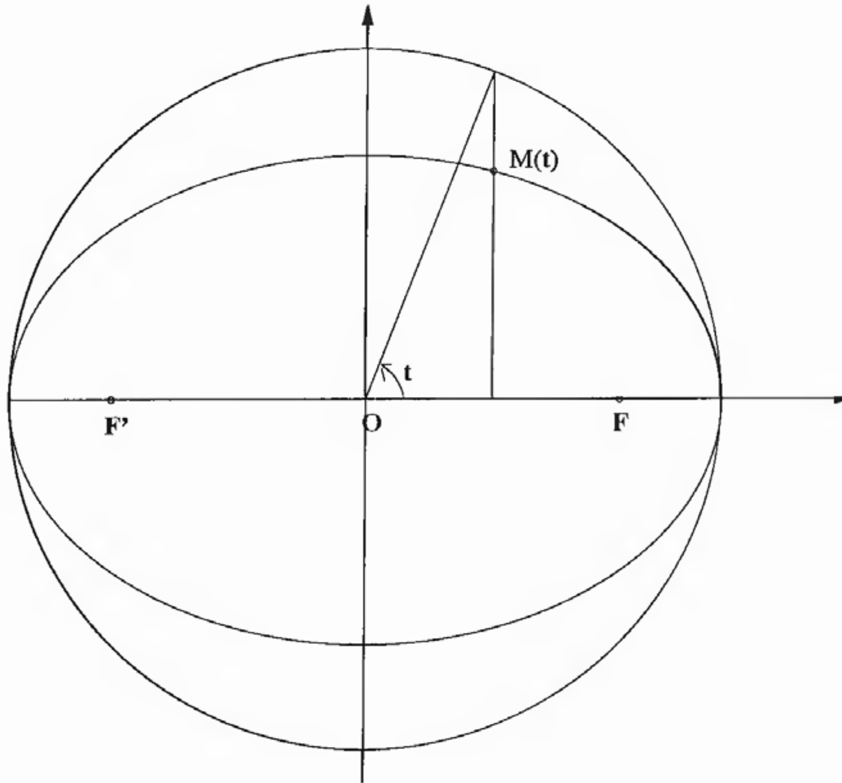



FIG. 6.1: *L'angle excentrique.*

**Propriété de la tangente.**

**Théorème 6. 3.** *Soit $\mathcal{E}$ une ellipse de foyers $F, F'$. Pour tout $M \in \mathcal{E}$, la normale et la tangente à $\mathcal{E}$ au point $M$ sont respectivement les bissectrices intérieure et extérieure de l'angle* $(\widehat{\mathbf{MF}, \mathbf{MF'}})$ *(voir ex. 7.2.).*

**Démonstration.**

Utilisons le paramétrage $t \mapsto \mathbf{M}(t)$ (6.4) de l'ellipse $\mathcal{E}$. Le vecteur $\frac{d\mathbf{OM}}{dt}$ est un vecteur directeur de la tangente à $\mathcal{E}$.

Soient $\mathbf{u} = \frac{\mathbf{FM}}{\|\mathbf{FM}\|}, \mathbf{v} = \frac{\mathbf{F'M}}{\|\mathbf{F'M}\|}$. Alors $\mathbf{u}+\mathbf{v}$ est un vecteur directeur de la bissectrice

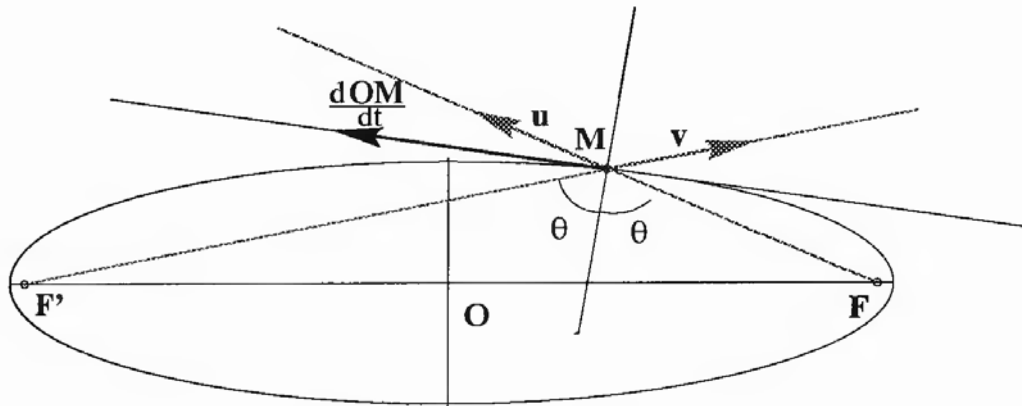



FIG. 6.2: *La propriété de la tangente pour l'ellipse.*

intérieure de l'angle $(\mathbf{MF},\mathbf{MF'})$ (voir Fig. 6.2). Il suffit donc de montrer que

$$(\mathbf{u}+\mathbf{v}|\frac{d\mathbf{OM}}{dt}) = 0. \tag{6.5}$$

L'équation bifocale $\|\mathbf{FM}\| + \|\mathbf{F'M}\| = 2a$ de $\mathcal{E}$ donne par dérivation

$$\frac{d\|\mathbf{FM}\|}{dt} + \frac{d\|\mathbf{F'M}\|}{dt} = 0. \tag{6.6}$$

Comme $\mathbf{FM} = \begin{pmatrix} x-c \\ y \end{pmatrix}$ où $c$ est la distance focale et $x = a\cos t, y = b\sin t$ sont les coordonnées du point $\mathbf{M}$, on a $\|\mathbf{FM}\| = \sqrt{(x-c)^2+y^2}$, donc

$$\frac{d\|\mathbf{FM}\|}{dt} = \frac{(x-c)\frac{dx}{dt} + y\frac{dy}{dt}}{\sqrt{(x-c)^2+y^2}} = \frac{(\mathbf{FM}|\frac{d\mathbf{OM}}{dt})}{\|\mathbf{FM}\|} = (\mathbf{u}|\frac{d\mathbf{OM}}{dt}).$$

On a de même $\|\mathbf{F'M}\| = \sqrt{(x+c)^2+y^2}$, donc

$$\frac{d\|\mathbf{F'M}\|}{dt} = \frac{(x+c)\frac{dx}{dt} + y\frac{dy}{dt}}{\sqrt{(x+c)^2+y^2}} = \frac{(\mathbf{F'M}|\frac{d\mathbf{OM}}{dt})}{\|\mathbf{F'M}\|} = (\mathbf{v}|\frac{d\mathbf{OM}}{dt}).$$

L'équation (6.6) donne alors l'équation (6.5). □

**Application.** Un rayon lumineux issu de l'un des foyers est réfléchi en un rayon passant par l'autre foyer.

### 6.1.2 Hyperbole.

**Définition bifocale.**

Soit $E$ le plan affine euclidien, $F, F'$ deux points distincts de $E$, $O$ le milieu du segment $[F, F']$, et $c = \|\mathbf{OF}\| > 0$.

Soit $\mathcal{H} = \{M \in E;\ |\,\|\mathbf{MF}\| - \|\mathbf{MF'}\|\,| = 2a\}$. Si $\mathcal{H} \neq \varnothing$ et $M \in \mathcal{H}$, on a d'après l'inégalité triangulaire $|\,\|\mathbf{MF}\| - \|\mathbf{MF'}\|\,| \leqslant \|\mathbf{FF'}\|$, *i.e.* $a \leqslant c$. Donc si $a > c$, $\mathcal{H} = \varnothing$. Si $a = c$ on a $\mathcal{H} = \{$ droite $FF'\} \setminus \{$ segment ouvert $]F, F'[\}$, si $0 < a < c$ il y a deux points du segment $]F, F'[$ qui appartiennent à $\mathcal{H}$, et si $a = 0$, $\mathcal{H}$ est la médiatrice du segment $[F, F']$.

**Définition 6. 4.** *Soit $E$ le plan affine euclidien, $F, F'$ deux points distincts de $E$, $O$ le milieu du segment $[F, F']$, $c = \|\mathbf{OF}\| > 0$, et $a < c$. On appelle hyperbole de foyers $F, F'$, de grand axe $a$ et de distance focale $c$ l'ensemble*

$$\mathcal{H} = \{M \in E;\ |\,\|\mathbf{MF}\| - \|\mathbf{MF'}\|\,| = 2a\}.$$

Comme dans le cas de l'ellipse, on prend comme origine du plan le point $O$ et on utilise un repère orthonormé affine $(O, (\mathbf{e_1}, \mathbf{e_2}))$ avec $\mathbf{e_1} = \frac{\mathbf{OF}}{\|\mathbf{OF}\|}$, et $\mathbf{e_2}$ vecteur déduit de $\mathbf{e_1}$ par rotation d'angle $\frac{\pi}{2}$. Nous identifions aussi le point $M$ au vecteur $\mathbf{OM}$ et notons $M = \begin{pmatrix} x \\ y \end{pmatrix}$.

L'axe focal est la droite $FF'$, de vecteur directeur $\mathbf{e_1}$ (*axe des $x$*), l'axe non focal est la droite perpendiculaire en $O$, de vecteur directeur $\mathbf{e_2}$ (*axe des $y$*). On voit comme pour l'ellipse que $O$ est centre de symétrie de $\mathcal{H}$. Les axes focal et non focal sont des axes de symétrie de $\mathcal{H}$. Il n'y a pas de point de l'axe non focal appartenant à $\mathcal{H}$. Comme pour l'ellipse, il y a deux points $A, A'$ symétriques par rapport à $O$ sur l'axe focal appartenant à $\mathcal{E}$ : ce sont les deux points d'abscisse $\pm a$.

On appelle *petit axe* le nombre $b > 0$ défini par $c^2 = a^2 + b^2$.

**Équation cartésienne.**

**Théorème 6. 4.** *L'équation cartésienne de l'hyperbole $\mathcal{H}$ de grand axe $a$ et petit axe $b$ est dans le repère orthonormé $(O, \mathbf{e_1}, \mathbf{e_2})$ :*

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \tag{6.7}$$

**Démonstration.**

Dans la base orthonormée choisie, on a $\mathbf{OF} = \begin{pmatrix} c \\ 0 \end{pmatrix}$ et $\mathbf{OM} = \begin{pmatrix} x \\ y \end{pmatrix}$ donc $\mathbf{FM} = \begin{pmatrix} x-c \\ y \end{pmatrix}$ et $\|\mathbf{FM}\| = \sqrt{(x-c)^2+y^2}$. De même $\mathbf{F'M} = \begin{pmatrix} x+c \\ y \end{pmatrix}$ et $\|\mathbf{F'M}\| = \sqrt{(x+c)^2+y^2}$. Ainsi :

$$\begin{aligned} M \in \mathcal{H} &\Leftrightarrow \; |\,\|\mathbf{FM}\| - \|\mathbf{F'M}\|\,| = 2a \\ &\Leftrightarrow \; \|\mathbf{FM}\|^2 + \|\mathbf{F'M}\|^2 - 2\|\mathbf{FM}\|\|\mathbf{F'M}\| = 4a^2 \\ &\Leftrightarrow \; 2(x^2+y^2+c^2) - 2\|\mathbf{FM}\|\|\mathbf{F'M}\| = 4a^2 \\ &\Leftrightarrow \; \|\mathbf{FM}\|\|\mathbf{F'M}\| = -2a^2 + (x^2+y^2+c^2) \\ &\Leftrightarrow \; \begin{cases} \|\mathbf{FM}\|^2\|\mathbf{F'M}\|^2 = 4a^4 - 4a^2(x^2+y^2+c^2) + (x^2+y^2+c^2)^2 \\ x^2+y^2+c^2 \geqslant 2a^2. \end{cases} \end{aligned}$$

Mais

$$\begin{aligned} \|\mathbf{FM}\|^2\|\mathbf{F'M}\|^2 &= (x^2+y^2+c^2-2xc)(x^2+y^2+c^2+2xc) \\ &= (x^2+y^2+c^2)^2 - 4x^2c^2, \end{aligned}$$

donc

$$\begin{aligned} M \in \mathcal{H} &\Leftrightarrow \begin{cases} -4x^2c^2 = 4a^4 - 4a^2(x^2+y^2+c^2) \\ x^2+y^2+c^2 \geqslant 2a^2 \end{cases} \\ &\Leftrightarrow \begin{cases} x^2(a^2-c^2) + a^2y^2 = a^2(a^2-c^2) \\ x^2+y^2+c^2 \geqslant 2a^2 \end{cases} \\ &\Leftrightarrow \begin{cases} \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \\ x^2+y^2+c^2 \geqslant 2a^2 \end{cases} \\ &\Leftrightarrow \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \end{aligned}$$

puisque la condition $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ implique $x^2 - \frac{a^2}{b^2}y^2 = a^2$ donc $x^2 + y^2 = a^2 + \frac{a^2+b^2}{b^2}y^2 \geqslant a^2$ et $x^2+y^2+c^2 \geqslant a^2+c^2 > 2a^2$. □

**Paramétrage.**

Soit $M \in \mathcal{H}$, $\mathbf{OM} = \begin{pmatrix} x \\ y \end{pmatrix}$. Comme l'application $t \mapsto \sinh t = \frac{e^t - e^{-t}}{2}$ est une bijection strictement croissante de $\mathbb{R}$ sur lui-même, il existe $t \in \mathbb{R}$ unique tel que $y = b \sinh t$. On a alors $\frac{x^2}{a^2} = 1 + \frac{y^2}{b^2} = 1 + \sinh^2 t = \cosh^2 t$ (où $\cosh t = \frac{e^t + e^{-t}}{2}$) donc $x = \varepsilon a \cosh t$ avec $\varepsilon = \pm 1$. Pour $\varepsilon = \pm 1$ fixé, l'application

$$t \mapsto M(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

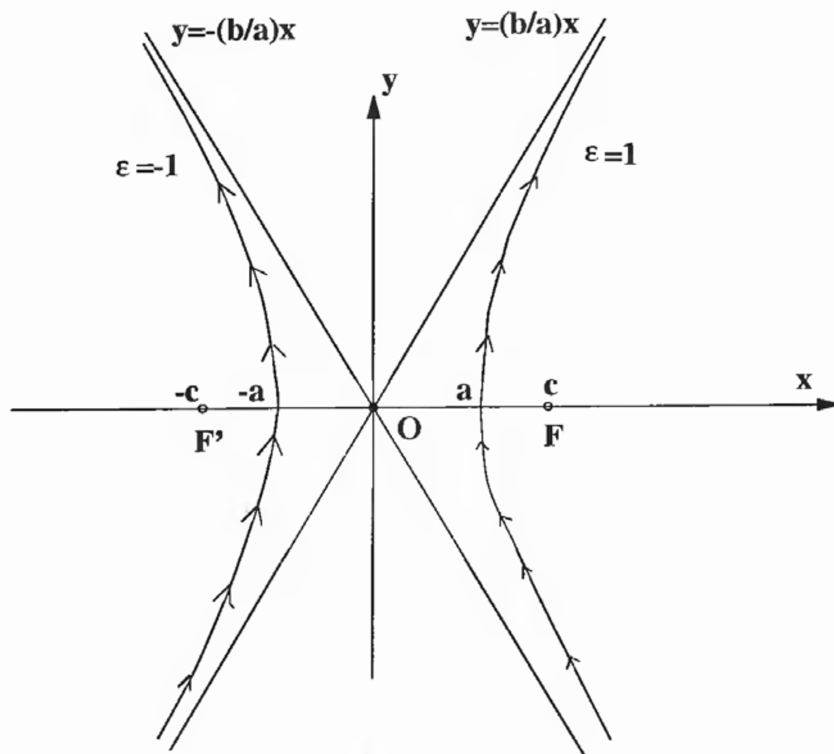



FIG. 6.3: *Branches et asymptotes de $\mathcal{H}$ avec sens de parcours de chaque branche dans le paramétrage (6.8).*

$$x(t) = \varepsilon a \cosh t, \quad y(t) = b \sinh t \tag{6.8}$$

est une bijection de $\mathbb{R}$ sur une partie de l'hyperbole $\mathcal{H}$, appelée *branche*. Il y a deux branches correspondant à $\varepsilon = 1$ et $\varepsilon = -1$ respectivement; les deux branches sont disjointes et $\mathcal{H}$ est leur réunion (voir Fig. 6.3).

Supposons $\varepsilon = 1$. Avec le paramétrage (6.8), on a

$$\begin{aligned} y - \frac{b}{a}x &= b(\sinh t - \cosh t) = -be^{-t} \to 0 \text{ quand } t \to +\infty, \\ y + \frac{b}{a}x &= b(\sinh t + \cosh t) = be^{t} \to 0 \text{ quand } t \to -\infty. \end{aligned}$$

La droite $y = \frac{b}{a}x$ est donc asymptote de la branche considérée lorsque $t \to +\infty$, et la droite $y = -\frac{b}{a}x$ est asymptote de la branche considérée lorsque $t \to -\infty$. Pour $\varepsilon = -1$, les résultats sont inversés : la droite $y = -\frac{b}{a}x$ est asymptote de la branche $\varepsilon = -1$ lorsque $t \to +\infty$, et la droite $y = \frac{b}{a}x$ est asymptote de la branche lorsque $t \to -\infty$. On constate donc que la courbe possède 2 asymptotes qui sont les droites $y = \pm\frac{b}{a}x$. Notons que l'équation de l'ensemble des 2 asymptotes s'écrit

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0. \tag{6.9}$$

**Définition par foyer-directrice.**

L'excentricité $e$ de l'hyperbole et la directrice associée à un foyer $F$ se définissent par les mêmes formules que pour l'ellipse (Définition 6.3). La seule différence est que

pour une hyperbole, on a $e > 1$. Comme dans le cas de l'ellipse, on a :

**Théorème 6. 5.** *Soit $\mathcal{H}$ une hyperbole, e son excentricité, F un foyer et $\mathcal{D}$ la directrice relative au foyer F.*

*Alors $\mathcal{H} = \{M \in E;\ \frac{\|\mathbf{MF}\|}{\|\mathbf{MH}\|} = e\}$, où le point H est le projeté orthogonal de M sur $\mathcal{D}$, i.e. $\|\mathbf{MH}\|$ est la distance du point M à $\mathcal{D}$.*

**Démonstration.**

La démonstration est identique à celle du théorème correspondant pour l'ellipse (Th. 6.2), à cela près que dans le cas de l'hyperbole on a $b^2 = c^2 - a^2$. On a $\|\mathbf{MF}\|^2 = (x-c)^2 + y^2$ et $\|\mathbf{MH}\|^2 = (x - \frac{a^2}{c})^2$.

L'équation $\|\mathbf{MF}\|^2 = e^2\|\mathbf{MH}\|^2$ s'écrit donc

$$(x-c)^2 + y^2 = \frac{c^2}{a^2}(x - \frac{a^2}{c})^2,$$

soit en développant :

$$x^2(a^2 - c^2) + y^2a^2 = a^2(a^2 - c^2)$$

ou encore

$$x^2(c^2 - a^2) - y^2a^2 = a^2(c^2 - a^2).$$

C'est l'équation (6.7) de l'hyperbole, d'où le résultat. □

**Propriété de la tangente.**

**Théorème 6. 6.** *Soit $\mathcal{H}$ une hyperbole de foyers $F, F'$. Pour tout $M \in \mathcal{H}$, la tangente et la normale à $\mathcal{H}$ au point M sont les bissectrices intérieure et extérieure de l'angle $(\widehat{\mathbf{MF}, \mathbf{MF'}})$ (voir ex. 7.2).*

**Démonstration.**

Utilisons le paramétrage $t \mapsto \mathrm{M}(t)$ (6.8) de l'hyperbole $\mathcal{H}$. Le vecteur $\frac{d\mathbf{OM}}{dt}$ est un vecteur directeur de la tangente à $\mathcal{E}$.

Soient $\mathbf{u} = \frac{\mathbf{FM}}{\|\mathbf{FM}\|}, \mathbf{v} = \frac{\mathbf{F'M}}{\|\mathbf{F'M}\|}$. Alors $\mathbf{u}+\mathbf{v}$ est un vecteur directeur de la bissectrice intérieure de l'angle $(\widehat{\mathbf{MF}, \mathbf{MF'}})$ (voir Fig.6.4). La bissectrice extérieure est perpendiculaire à la bissectrice intérieure. Or le vecteur $\mathbf{u} - \mathbf{v}$ est orthogonal au vecteur $\mathbf{u}+\mathbf{v}$ puisque $(\mathbf{u}+\mathbf{v} \,|\, \mathbf{u}+\mathbf{v}) = \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2 = 0$. Donc $\mathbf{u}-\mathbf{v}$ est un vecteur directeur de la bissectrice extérieure. Il suffit donc de montrer que

$$(\mathbf{u} - \mathbf{v}|\frac{d\mathbf{OM}}{dt}) = 0. \tag{6.10}$$

L'équation bifocale de $\mathcal{H}$ est $|\,\|\mathbf{FM}\| - \|\mathbf{F'M}\|\,| = 2a$. Sur chacune des branches de $\mathcal{H}$, la différence $\|\mathbf{FM}\| - \|\mathbf{F'M}\|$ qui est une fonction continue du paramètre $t$ est constante, égale à $2a$ ou $-2a$ suivant la branche. Par dérivation on obtiendra donc

$$\frac{d\|\mathbf{FM}\|}{dt} - \frac{d\|\mathbf{F'M}\|}{dt} = 0. \tag{6.11}$$

La suite du raisonnement est identique à celui fait pour l'ellipse. Comme $\mathbf{FM} = \begin{pmatrix} x-c \\ y \end{pmatrix}$ où $c$ est la distance focale et $x = \varepsilon a \cosh t, y = b \sinh t$ sont les coordonnées du point $\mathbf{M}$, on a $\|\mathbf{FM}\| = \sqrt{(x-c)^2 + y^2}$, donc

$$\frac{d\|\mathbf{FM}\|}{dt} = \frac{(x-c)\frac{dx}{dt} + y\frac{dy}{dt}}{\sqrt{(x-c)^2+y^2}} = \frac{(\mathbf{FM}|\frac{d\mathbf{OM}}{dt})}{\|\mathbf{FM}\|} = (\mathbf{u}|\frac{d\mathbf{OM}}{dt}).$$

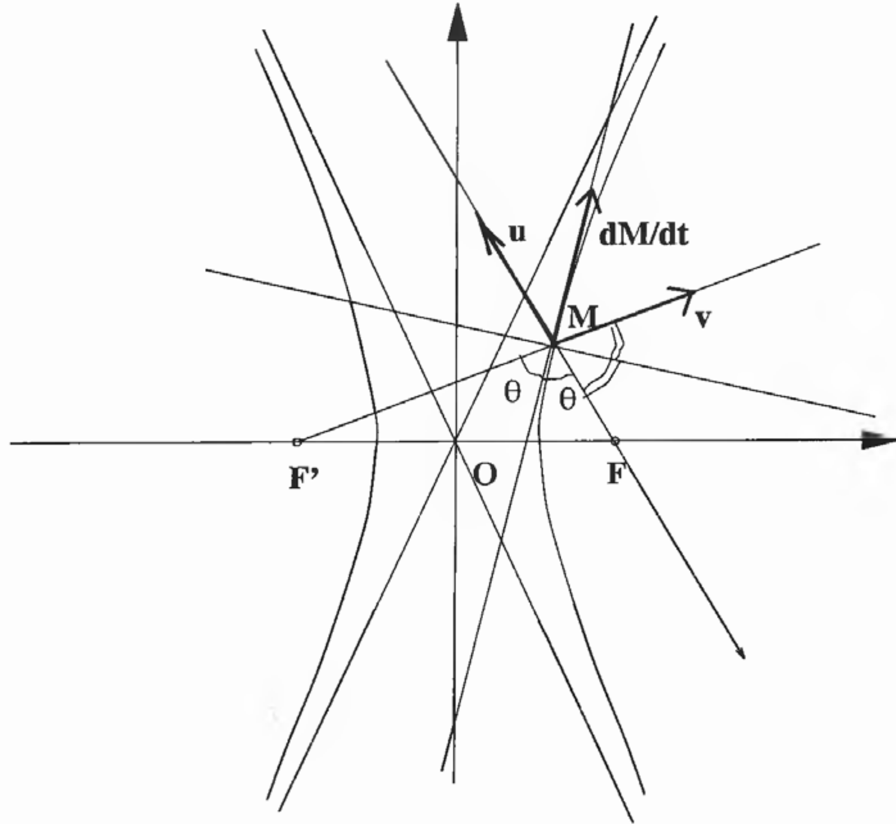



FIG. 6.4: *La propriété de la tangente pour l'hyperbole.*

On a de même $\|\mathbf{F'M}\| = \sqrt{(x+c)^2+y^2}$, donc

$$\frac{d\|\mathbf{F'M}\|}{dt} = \frac{(x+c)\frac{dx}{dt} + y\frac{dy}{dt}}{\sqrt{(x+c)^2+y^2}} = \frac{(\mathbf{F'M}|\frac{d\mathbf{OM}}{dt})}{\|\mathbf{F'M}\|} = (\mathbf{v}|\frac{d\mathbf{OM}}{dt}).$$

L'équation (6.11) donne alors l'équation (6.10). □

**Application.** Un rayon lumineux issu de l'un des foyers est réfléchi en un rayon ayant pour *source virtuelle* l'autre foyer.

### 6.1.3 Parabole.

**Définition, équation cartésienne.**

**Définition 6. 5.** *Soit $\mathcal{D}$ une droite et $F$ un point, $F \notin \mathcal{D}$. On appelle parabole $\mathcal{P}$ de foyer $F$ et directrice $\mathcal{D}$ l'ensemble des points $M$ du plan équidistants de $F$ et $\mathcal{D}$ :*

$$\mathcal{P} = \{M \in E;\ \frac{\|\mathbf{MF}\|}{\|\mathbf{MH}\|} = 1\},$$

*où le point $H$ est le projeté orthogonal de $M$ sur $\mathcal{D}$, i.e. $\|\mathbf{MH}\|$ est la distance du point $M$ à $\mathcal{D}$ (voir Fig. 6.5).*

L'excentricité $e$ d'une parabole est définie comme étant égale à 1 :

$$e = 1.$$

Soit $K$ le projeté orthogonal du foyer $F$ sur la directrice $\mathcal{D}$. Le milieu $O$ du segment $[K, F]$ est un point de la parabole. On prend comme origine du plan le point $O$ et on utilise un repère orthonormé affine $(O, (\mathbf{e_1}, \mathbf{e_2}))$ avec $\mathbf{e_1} = \frac{\mathbf{OF}}{\|\mathbf{OF}\|}$, et $\mathbf{e_2}$ vecteur déduit de $\mathbf{e_1}$ par rotation d'angle $\frac{\pi}{2}$. L'axe focal est la droite $OF$, de vecteur directeur $\mathbf{e_1}$ (*axe des $x$*). C'est un axe de symétrie de $\mathcal{P}$.

**Paramètre d'une conique.**

**Définition 6. 6.** *Etant donnée une conique $\mathcal{C}$ (ellipse, hyperbole ou parabole) et $F$ un foyer de $\mathcal{C}$, on appelle paramètre de $\mathcal{C}$ la distance $p$ à $F$ des points de $\mathcal{C}$ dont le projeté orthogonal sur l'axe focal est $F$.*

Pour une conique à centre, *i.e.* une ellipse ou une hyperbole, l'axe non focal étant un axe de symétrie, le paramètre ne dépend pas du foyer utilisé dans la définition.

Pour une parabole, le paramètre est $p = d$ en posant $d = \|\mathbf{KF}\|$ distance du foyer à la directrice. Pour une conique quelconque d'excentricité $e > 0$, le paramètre est

$$p = ed, \tag{6.12}$$

en notant encore $d$ la distance de $F$ à la directrice associée. On a aussi :

**Proposition 6. 1.** *Soit $\mathcal{C}$ une conique à centre, i.e. une ellipse ou une hyperbole, de grand axe $a$, petit axe $b$. Le paramètre $p$ de $\mathcal{C}$ est :*

$$p = \frac{b^2}{a}. \tag{6.13}$$

**Démonstration.**

Soit $F$ un foyer, $\mathcal{D}$ la directrice associée et $M$ un point de $\mathcal{C}$ dont le projeté orthogonal sur l'axe focal est $F$. On utilise le repère orthonormé direct habituel $(O, (\mathbf{e_1}, \mathbf{e_2}))$ avec $O$ centre de $\mathcal{C}$ et $\mathbf{e_1} = \frac{\mathbf{OF}}{\|\mathbf{OF}\|}$. Par définition de $p$, on a $p = \|\mathbf{MF}\|$. D'après la définition par foyer-directrice de $\mathcal{C}$,

$$p = \|\mathbf{MH}\| e \tag{6.14}$$

où $H$ est le projeté orthogonal de $M$ sur $\mathcal{D}$ et $e$ l'excentricité de $\mathcal{C}$. Or $\|\mathbf{MH}\| = \|\mathbf{FK}\|$, $K$ désignant le pied de la directrice sur l'axe focal, et $\mathbf{FK} = \mathbf{OK} - \mathbf{OF} = \frac{a^2}{c}\mathbf{e_1} - c\mathbf{e_1}, = \frac{a^2-c^2}{c}\mathbf{e_1}$, $c$ étant la distance focale. Donc $\|\mathbf{MH}\| = \frac{|a^2-c^2|}{c} = \frac{b^2}{c}$, et par report dans (6.14), on obtient $p = \frac{b^2}{a}$ puisque $e = \frac{c}{a}$. □

**Équation cartésienne de la parabole.**

**Théorème 6. 7.** *L'équation cartésienne de la parabole $\mathcal{P}$ foyer $F$ et directrice $\mathcal{D}$ est dans le repère orthonormé $(O, (\mathbf{e_1}, \mathbf{e_2}))$ :*

$$y^2 = 2px \tag{6.15}$$

*où $p$ est le paramètre.*

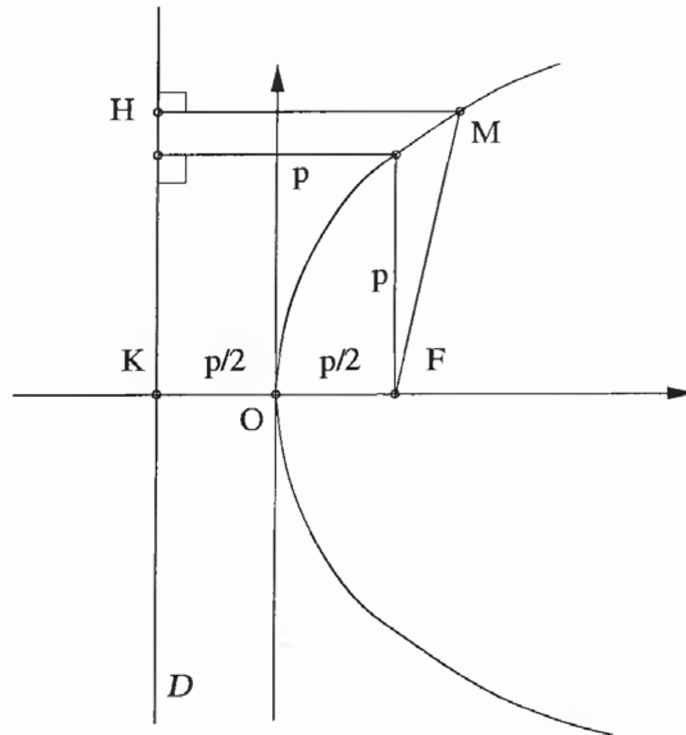



FIG. 6.5: *La parabole de foyer $F$ et directrice $\mathcal{D}$.*

**Démonstration.**

Dans la base orthonormée choisie, on a $\mathbf{OF} = \begin{pmatrix} \frac{p}{2} \\ 0 \end{pmatrix}$ et $\mathbf{OM} = \begin{pmatrix} x \\ y \end{pmatrix}$ donc $\mathbf{FM} = \begin{pmatrix} x - \frac{p}{2} \\ y \end{pmatrix}$ et $\|\mathbf{FM}\| = \sqrt{(x-\frac{p}{2})^2 + y^2}$. De même $\mathbf{HM} = \begin{pmatrix} x + \frac{p}{2} \\ 0 \end{pmatrix}$ et $\|\mathbf{F'M}\| = \sqrt{(x+\frac{p}{2})^2 + y^2}$. Ainsi :

$$\begin{aligned} M \in \mathcal{P} &\Leftrightarrow \|\mathbf{FM}\| = \|\mathbf{MH}\| \\ &\Leftrightarrow (x - \frac{p}{2})^2 + y^2 = (x + \frac{p}{2})^2 \\ &\Leftrightarrow y^2 = 2px. \end{aligned}$$

□

**Paramétrage.**

Soit $M \in \mathcal{P}$, $\mathbf{OM} = \begin{pmatrix} x \\ y \end{pmatrix}$. D'après l'équation (6.15), l'application

$$t \mapsto M(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

$$x(t) = \frac{t^2}{2p}, \quad y(t) = t \tag{6.16}$$

est une bijection de $\mathbb{R}$ sur $\mathcal{P}$.

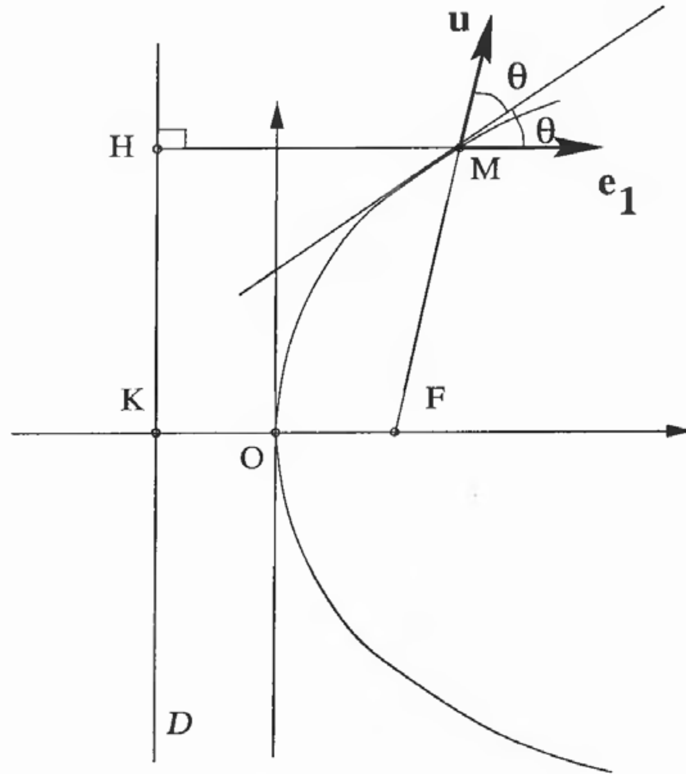



FIG. 6.6: *La propriété de la tangente pour la parabole.*

**Propriété de la tangente.**

**Théorème 6. 8.** *Soit $\mathcal{P}$ une parabole de foyer $F$ et de directrice $\mathcal{D}$. Pour tout $M \in \mathcal{P}$, la tangente et la normale à $\mathcal{P}$ au point $M$ sont les bissectrices intérieure et extérieure de l'angle* $(\widehat{\mathbf{MF}, \mathbf{MH}})$, *où $H$ désigne le projeté orthogonal de $M$ sur $\mathcal{D}$ (voir ex. 7.2).*

**Démonstration.**

Utilisons le paramétrage $t \mapsto \mathrm{M}(t)$ (6.16) de la parabole $\mathcal{P}$. Le vecteur $\frac{d\mathbf{OM}}{dt}$ est un vecteur directeur de la tangente à $\mathcal{P}$.

Soit $\mathbf{u} = \frac{\mathbf{FM}}{\|\mathbf{FM}\|}$. Le vecteur $\frac{\mathbf{HM}}{\|\mathbf{HM}\|}$ n'est autre que le vecteur $\mathbf{e_1}$. Donc $\mathbf{u} + \mathbf{e_1}$ est un vecteur directeur de la bissectrice intérieure de l'angle $(\widehat{\mathbf{MF}, \mathbf{MH}})$ (voir Fig. 6.6). La bissectrice extérieure est perpendiculaire à la bissectrice intérieure. Or le vecteur $\mathbf{u} - \mathbf{e_1}$ est orthogonal au vecteur $\mathbf{u} + \mathbf{e_1}$. Donc $\mathbf{u} - \mathbf{e_1}$ est un vecteur directeur de la bissectrice extérieure. Il suffit ainsi de montrer que

$$(\mathbf{u} - \mathbf{e_1} | \frac{d\mathbf{OM}}{dt}) = 0. \tag{6.17}$$

L'équation de $\mathcal{P}$ est $\|\mathbf{MF}\| = \|\mathbf{MH}\|$. Comme $\mathbf{FM} = \begin{pmatrix} x - \frac{p}{2} \\ y \end{pmatrix}$, et $\mathbf{HM} = \begin{pmatrix} x + \frac{p}{2} \\ 0 \end{pmatrix}$, où $(x, y)$ sont les coordonnées du point $\mathbf{M}$, cette équation s'écrit :

$$\sqrt{(x - \frac{p}{2})^2 + y^2} = x + \frac{p}{2}.$$

D'où par dérivation

$$\frac{(x-\frac{p}{2})\frac{dx}{dt}+y\frac{dy}{dt}}{\sqrt{(x-\frac{p}{2})^2+y^2}}=\frac{dx}{dt}.$$

Cela s'écrit

$$(\mathbf{u}|\frac{d\mathbf{OM}}{dt})=(\mathbf{e_1}|\frac{d\mathbf{OM}}{dt}).$$

D'où l'équation (6.17). □

**Application.** Un rayon lumineux issu du foyer est réfléchi en un rayon horizontal et réciproquement (miroir parabolique, antenne parabolique).

### 6.1.4 Équation d'une conique en coordonnées polaires.

**Rappels.**

Rappelons que, étant donnés un *pôle* $O$ et un *axe polaire orienté* $Ox$ de vecteur directeur normé $\mathbf{i}$, on appelle *système de coordonnées polaires* d'un point $M$ tout couple $(r,\theta)\in\mathbb{R}^2$ tel que $\mathbf{OM}=r\,\mathbf{u}(\theta)$ où $\mathbf{u}(\theta)=\cos\theta\ \mathbf{i}+\sin\theta\ \mathbf{j}$, avec $\mathbf{j}$ vecteur directement perpendiculaire à $\mathbf{i}$.

Si $(r,\theta)$ est un système de coordonnées polaires de $M$, il en est de même de $((-1)^p r,\theta+p\pi)$, $\forall p\in\mathbb{Z}$.

Etant donnée une courbe $\Gamma$ du plan, une *équation en coordonnées polaires de* $\Gamma$ est une équation $r=f(\theta)$ où $f$ est une fonction, définie sur une partie $I$ de $\mathbb{R}$, telle que

$$\Gamma=\{M\,;\exists\theta\in I\quad \mathbf{OM}=f(\theta)\mathbf{u}(\theta)\}$$

*i.e.* $\Gamma$ est l'ensemble de points $M$ du plan pour lesquels il existe $\theta\in I$ tel que $(f(\theta),\theta)$ soit un système de coordonnées polaires de $M$.

Si $r=f(\theta)$ est une équation en coordonnées polaires de $\Gamma$, et si $M$ est un point de $\Gamma$, il existe $\theta\in I$ tel que le couple $(f(\theta),\theta)$ soit un système de coordonnées polaires de $M$. Pour tout $p\in\mathbb{Z}$, $((-1)^p f(\theta),\theta+p\pi)$ est alors un autre système de coordonnées polaires de ce même point $M$. Si l'on pose $\varphi=\theta+p\pi$, ce dernier système est $((-1)^p f(\varphi-p\pi),\varphi)$.

Il en résulte que, pour tout $p\in\mathbb{Z}$, l'équation $r=(-1)^p\,f(\theta-p\pi)$ est une autre équation en coordonnées polaires de la même courbe $\Gamma$. En particulier en prenant $p=1$, $r=-f(\theta-\pi)$ est une autre équation en coordonnées polaires de $\Gamma$.

**Équation d'une conique.**

Soit $\Gamma$ une conique qui n'est pas un cercle, $F$ un foyer et $\mathcal{D}$ la directrice associée. On prend comme pôle $O=F$ et comme axe polaire orienté $\mathbf{OK}$, $K$ étant le projeté orthogonal de $F$ sur $\mathcal{D}$. Comme plus haut, $\mathbf{i}$ est le vecteur unitaire de l'axe polaire orienté, et $\mathbf{j}$ le vecteur directement perpendiculaire à $\mathbf{i}$.

**Théorème 6. 9.** *Dans les conditions ci-dessus, une équation en coordonnées polaires d'une conique* $\Gamma$ *d'excentricité* $e$ *est*

$$r=\frac{p}{1+e\cos\theta}\tag{6.18}$$

*où* $p=ed$ *est le paramètre de la conique* $\Gamma$, *d étant la distance du foyer à la directrice associée.*

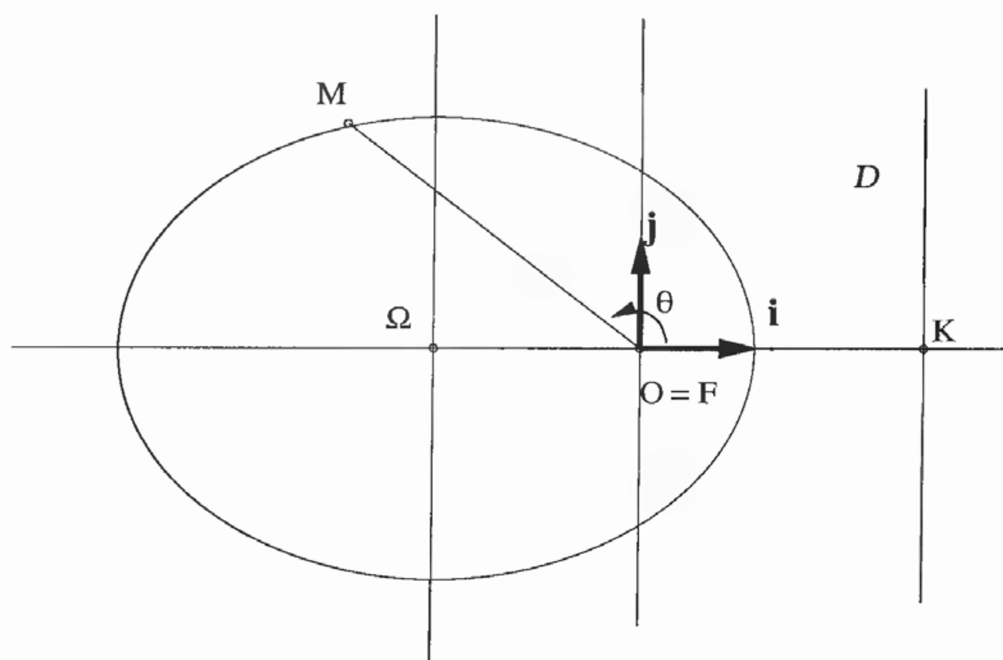



FIG. 6.7: *Équation en coordonnées polaires de l'ellipse.*

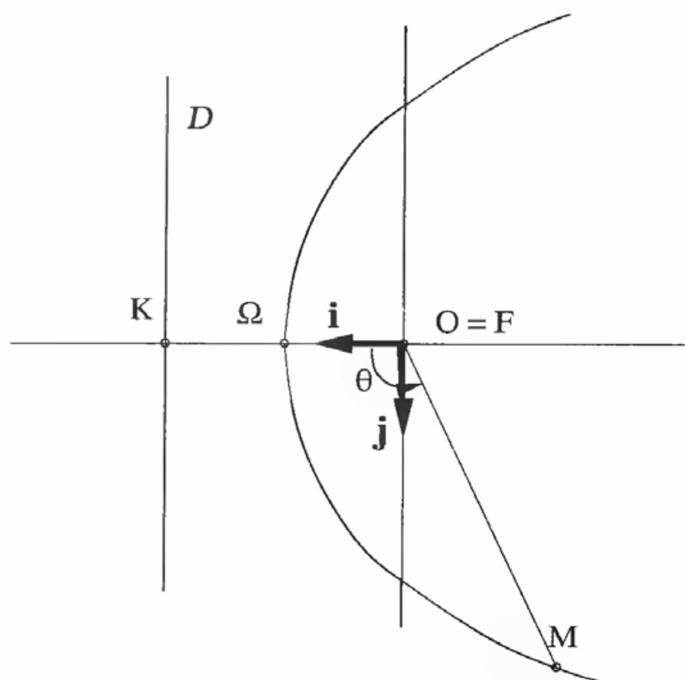



FIG. 6.8: *Équation en coordonnées polaires de la parabole.*

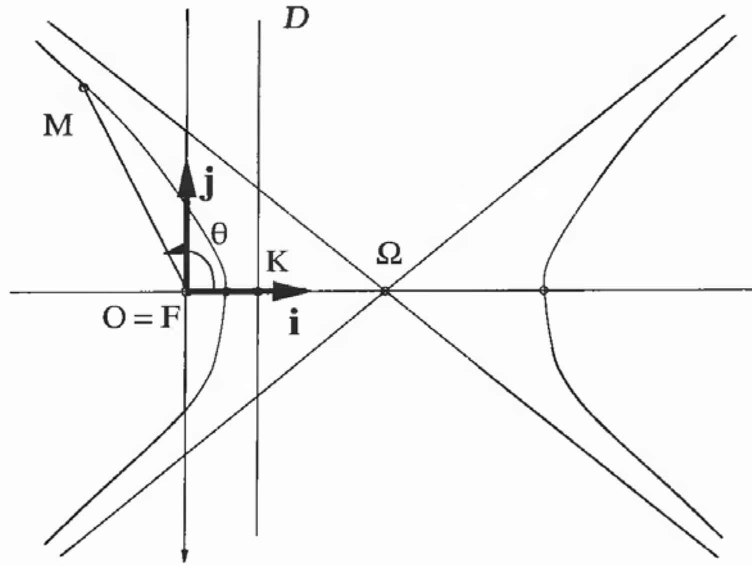



FIG. 6.9: *Équation en coordonnées polaires de l'hyperbole.*

Soit $f$ un polynôme de degré 2 en $x, y$ :

$$f(x,y) = \alpha x^2 + 2\beta xy + \gamma y^2 + \delta x + \varepsilon y + \eta$$

avec $\alpha, \beta, \gamma, \delta, \varepsilon, \eta \in \mathbb{R}$ et $\alpha, \beta, \gamma$ *non tous trois nuls.* Soit

$$\mathcal{C} = \{M\,;\mathbf{OM} = \begin{pmatrix} x \\ y \end{pmatrix};\ f(x,y) = 0\}. \tag{6.21}$$

Nous allons étudier l'ensemble $\mathcal{C}$.

Soit

$$q(\mathbf{OM}) = \alpha x^2 + 2\beta xy + \gamma y^2$$

la partie homogène de degré 2 de $f$. On a

$$q(\mathbf{OM}) = {}^t\begin{pmatrix} x \\ y \end{pmatrix} A \begin{pmatrix} x \\ y \end{pmatrix}$$

avec

$$A = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}.$$

La matrice $A$ est symétrique réelle, donc *orthodiagonalisable.* Soit $(\mathbf{f_1}, \mathbf{f_2})$ une base orthonormée de $\mathbb{R}^2$ diagonalisant $A$ , *i.e.* formée de vecteurs propres de $A$, et $P$ la matrice de passage de la base orthonormée $(\mathbf{e_1}, \mathbf{e_2})$ à la base $(\mathbf{f_1}, \mathbf{f_2})$. La matrice $P$ est *orthogonale* et l'on a :

$${}^tP\,AP = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \tag{6.22}$$

avec $\lambda, \mu \in \mathbb{R}$. Les deux nombres $\lambda, \mu$ ne sont pas simultanément nuls puisque la matrice $A$ est non nulle. Si $(X, Y)$ désignent les coordonnées de $\mathbf{OM}$ dans la base $(\mathbf{f_1}, \mathbf{f_2})$, on a

$$\begin{pmatrix} x \\ y \end{pmatrix} = P \begin{pmatrix} X \\ Y \end{pmatrix} \tag{6.23}$$

donc

$$\begin{aligned} q(\mathbf{OM}) &= {}^t\begin{pmatrix} x \\ y \end{pmatrix} A \begin{pmatrix} x \\ y \end{pmatrix} \\ &= {}^t\left(P\begin{pmatrix} X \\ Y \end{pmatrix}\right) AP \begin{pmatrix} X \\ Y \end{pmatrix} \\ &= {}^t\begin{pmatrix} X \\ Y \end{pmatrix} {}^tP\,AP \begin{pmatrix} X \\ Y \end{pmatrix} \\ &= {}^t\begin{pmatrix} X \\ Y \end{pmatrix} \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} \\ &= \lambda X^2 + \mu Y^2. \end{aligned}$$

Par ailleurs, $\delta x + \varepsilon y$ est de la forme $2\varrho X + 2\sigma Y$ d'après (6.23). Finalement

$$f(x, y) = \lambda X^2 + \mu Y^2 + 2\varrho X + 2\sigma Y + \eta,$$

*i.e.* $\mathcal{C}$ est l'ensemble des points du plan dont les coordonnées $(X, Y)$ par rapport au repère orthonormé $(O, (\mathbf{f_1}, \mathbf{f_2}))$ vérifient l'équation

$$\lambda X^2 + \mu Y^2 + 2\varrho X + 2\sigma Y + \eta = 0. \tag{6.24}$$

**1er cas :** l'un des deux nombres $\lambda, \mu$ est nul. On peut supposer $\lambda = 0, \mu \neq 0$, le raisonnement étant analogue si c'est $\mu$ qui est nul. L'équation (6.24) s'écrit alors

$$\mu(Y + \frac{\sigma}{\mu})^2 + 2\varrho X + \eta - \frac{\sigma^2}{\mu} = 0. \tag{6.25}$$

• Si $\varrho \neq 0$, (6.25) s'écrit :

$$\mu(Y + \frac{\sigma}{\mu})^2 = -2\varrho\,(X + \frac{\eta}{2\varrho} - \frac{\sigma^2}{2\varrho\mu}).$$

C'est l'équation d'une parabole dont le sommet $\Omega$ a pour coordonnées

$$\begin{cases} X_\Omega &= -\frac{\eta}{2\varrho} + \frac{\sigma^2}{2\varrho\mu} \\ Y_\Omega &= -\frac{\sigma}{\mu}. \end{cases}$$

Le paramètre est $p = \left|\frac{\varrho}{\mu}\right|$.

• Si $\varrho = 0$, (6.25) s'écrit :

$$\mu(Y + \frac{\sigma}{\mu})^2 = -\eta + \frac{\sigma^2}{\mu}.$$

$\mathcal{C}$ est donc l'ensemble vide si $\frac{\sigma^2}{\mu^2} < \frac{\eta}{\mu}$, la droite $Y = -\frac{\sigma}{\mu}$ si $\frac{\sigma^2}{\mu^2} = \frac{\eta}{\mu}$, et enfin l'ensemble des deux droites parallèles $Y + \frac{\sigma}{\mu} = \pm\sqrt{\frac{\sigma^2}{\mu^2} - \frac{\eta}{\mu}}$ si $\frac{\sigma^2}{\mu^2} > \frac{\eta}{\mu}$.

**2ème cas :** $\lambda\mu \neq 0$. L'équation (6.24) s'écrit alors

$$\lambda(X + \frac{\varrho}{\lambda})^2 + \mu(Y + \frac{\sigma}{\mu})^2 + \eta - \frac{\varrho^2}{\lambda} - \frac{\sigma^2}{\mu} = 0 \tag{6.26}$$

ou encore

$$\lambda(X - X_\Omega)^2 + \mu(Y - Y_\Omega)^2 + \tau = 0 \tag{6.27}$$

en notant $\Omega$ le point de coordonnées

$$\begin{cases} X_\Omega &= -\frac{\varrho}{\lambda} \\ Y_\Omega &= -\frac{\sigma}{\mu} \end{cases}$$

et

$$\tau = \eta - \frac{\varrho^2}{\lambda} - \frac{\sigma^2}{\mu}.$$

• Si $\tau = 0$, $\mathcal{C}$ est l'ensemble des deux droites $Y - Y_\Omega = \pm\sqrt{-\frac{\lambda}{\mu}}(X - X_\Omega)$ dans le cas $\lambda\mu < 0$, et $\mathcal{C}$ est réduit au point $\Omega$ si $\lambda\mu > 0$.

• Si $\tau \neq 0$, $\mathcal{C}$ est l'ensemble vide si $\lambda$ et $\mu$ sont tous deux du signe de $\tau$, une ellipse (éventuellement dégénérée en cercle) si $\lambda$ et $\mu$ sont tous deux du signe contraire de $\tau$, et une hyperbole si $\lambda$ et $\mu$ sont de signes contraires.

**Remarque.** Dans le cas où $\mathcal{C}$ est une ellipse ou une hyperbole, les coordonnées $(x_\Omega, y_\Omega)$ du centre $\Omega$ sont les solutions du système

$$\begin{cases} \frac{\partial f}{\partial x} &= 0 \\ \frac{\partial f}{\partial y} &= 0. \end{cases} \tag{6.28}$$

En effet d'après (6.27), le changement de variable (6.23) dans $f(x,y)$ a donné

$$f(x,y) = g(X,Y)$$

avec

$$g(X,Y) = \lambda(X - X_\Omega)^2 + \mu(Y - Y_\Omega)^2 + \tau\,. \tag{6.29}$$

D'après (6.29), $X_\Omega, Y_\Omega$ sont les solutions du système

$$\begin{cases} \frac{\partial g}{\partial X} &= 0 \\ \frac{\partial g}{\partial Y} &= 0\,. \end{cases} \tag{6.30}$$

Mais le changement de variable (6.23) donne par différentiation

$$\begin{aligned} \frac{\partial g}{\partial X} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial X} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial X} \\ &= \frac{\partial f}{\partial x}p_1^1 + \frac{\partial f}{\partial y}p_1^2 \\ \frac{\partial g}{\partial Y} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial Y} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial Y} \\ &= \frac{\partial f}{\partial x}p_2^1 + \frac{\partial f}{\partial y}p_2^2 \end{aligned}$$

*i.e.*

$$\begin{pmatrix} \frac{\partial g}{\partial X} \\ \frac{\partial g}{\partial Y} \end{pmatrix} = {}^tP \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} \tag{6.31}$$

en notant $p_i^j \quad 1 \leqslant i,j \leqslant 2$ les coefficients de $P$. Comme $P$ est inversible, les systèmes (6.30) et (6.28) sont équivalents et donc $x_\Omega, x_\Omega$ sont les solutions du système (6.28).

## 6.2 Cercles.

Soit $E$ le plan euclidien muni d'un repère orthonormé $(O, (\mathbf{e_1}, \mathbf{e_2}))$ et notons $(x, y)$ les coordonnées d'un point $M$ :

$$\mathbf{OM} = \begin{pmatrix} x \\ y \end{pmatrix} = x\mathbf{e_1} + y\mathbf{e_2}.$$

### 6.2.1 Équation cartésienne.

Le *cercle* de centre le point $I$ de coordonnées $(a, b)$

$$\mathbf{OI} = \begin{pmatrix} a \\ b \end{pmatrix},$$

et de rayon $R > 0$ est l'ensemble $\mathcal{C}(I, R)$ des points $M$ tels que $\|\mathbf{IM}\|^2 = R^2$. Son équation est

$$(x-a)^2 + (y-b)^2 = R^2.$$

**Proposition 6. 2.** *Soient $a, b, c \in \mathbb{R}$. L'ensemble des points $M$ du plan dont les coordonnées $(x, y)$ vérifient l'équation*

$$x^2 + y^2 - 2ax - 2by + c = 0 \tag{6.32}$$

*est un cercle de rayon $R = \sqrt{a^2 + b^2 - c}$ et de centre le point $I$ de coordonnées $(a, b)$ si $a^2 + b^2 \geqslant c$, et est l'ensemble vide sinon.*

**Démonstration.**
L'équation (6.32) s'écrit en effet :

$$(x-a)^2 + (y-b)^2 = a^2 + b^2 - c.$$

□

### 6.2.2 Puissance d'un point par rapport à un cercle.

Soit $\Gamma$ le cercle d'équation (6.32) avec $a^2 + b^2 \geqslant c$, $I$ son centre et $R$ son rayon. Pour tout point $M$ de coordonnées $(x, y)$, notons

$$\mathcal{P}_\Gamma(M) = x^2 + y^2 - 2ax - 2by + c. \tag{6.33}$$

**Lemme 6. 1.** *Pour tout point $M$ du plan on a $\mathcal{P}_\Gamma(M) = \|\mathbf{IM}\|^2 - R^2$. En particulier, $\mathcal{P}_\Gamma(M)$ ne dépend pas du repère orthonormé utilisé, et*

$$\mathcal{P}_\Gamma(M) = \begin{cases} 0 & si\ M \in \Gamma \\ > 0 & si\ \|\mathbf{IM}\| > R \\ < 0 & si\ \|\mathbf{IM}\| < R. \end{cases}$$

**Démonstration.**
On a en effet

$$\|\mathbf{IM}\|^2 - R^2 = (x-a)^2 + (y-b)^2 - (a^2 + b^2 - c) = x^2 + y^2 - 2ax - 2by + c.$$

□

**Définition 6. 7.** *La fonction $\mathcal{P}_\Gamma : M \mapsto \mathcal{P}_\Gamma(M)$ s'appelle la puissance par rapport à $\Gamma$, et $\mathcal{P}_\Gamma(M)$ la puissance de $M$ par rapport à $\Gamma$.*

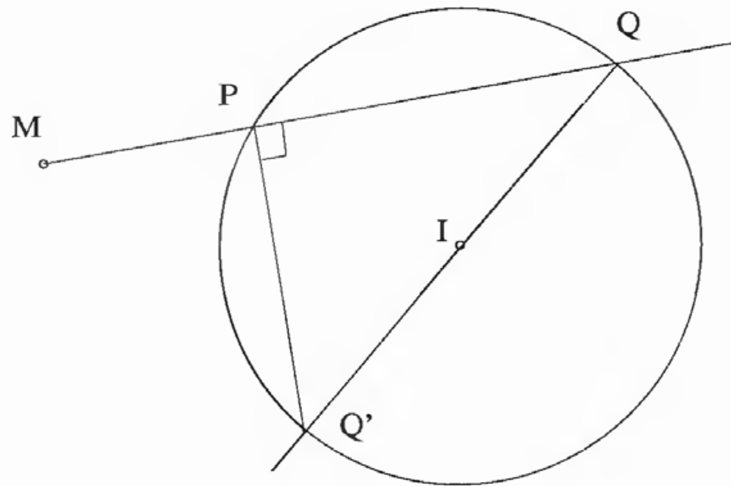



FIG. 6.10: *Puissance d'un point par rapport à un cercle.*

**Proposition 6. 3.** *Soit $\Gamma = \mathcal{C}(I, R)$ un cercle de centre $I$ et de rayon $R$. Si une droite issue d'un point $M$ rencontre $\Gamma$ en deux points $P, Q$ (éventuellement confondus), on a :*

$$\mathcal{P}_\Gamma(M) = (\mathbf{MP}|\mathbf{MQ}).$$

**Démonstration.**
Soit $Q'$ le point diamétralement opposé à $Q$ (voir Fig. 6.10). On a $\mathbf{MP} = \mathbf{MQ'} + \mathbf{Q'P}$. Si $P \neq Q$, le triangle $QQ'P$ est rectangle en $P$, donc le vecteur $\mathbf{Q'P}$ est orthogonal au vecteur $\mathbf{MQ}$. Cette dernière propriété est encore vraie si $P = Q$. Donc dans les deux cas $(\mathbf{Q'P}|\mathbf{MQ}) = 0$. On en déduit :

$$\begin{aligned}
(\mathbf{MP}|\mathbf{MQ}) &= (\mathbf{MQ'}|\mathbf{MQ}) + (\mathbf{Q'P}|\mathbf{MQ}) \\
&= (\mathbf{MQ'}|\mathbf{MQ}) \\
&= (\mathbf{MI} + \mathbf{IQ'}|\mathbf{MI} + \mathbf{IQ}) \\
&= (\mathbf{MI} - \mathbf{IQ}|\mathbf{MI} + \mathbf{IQ}) \\
&= \|\mathbf{MI}\|^2 - R^2 \\
&= \mathcal{P}_\Gamma(M)\,.
\end{aligned}$$

□

### 6.2.3 Axe radical.

**Axe radical de 2 cercles.**

**Définition 6. 8.** *Soient $\Gamma, \Gamma'$ deux cercles dont les centres $I, I'$ sont distincts. On appelle axe radical de $\Gamma$ et $\Gamma'$ l'ensemble $\Delta$ des points $M$ du plan qui ont même puissance par rapport à $\Gamma$ et $\Gamma'$ :*

$$\Delta = \{M;\ \mathcal{P}_\Gamma(M) = \mathcal{P}_{\Gamma'}(M)\}.$$

**Lemme 6. 2.** *L'axe radical de deux cercles dont les centres sont distincts est une droite perpendiculaire à la droite joignant les centres.*

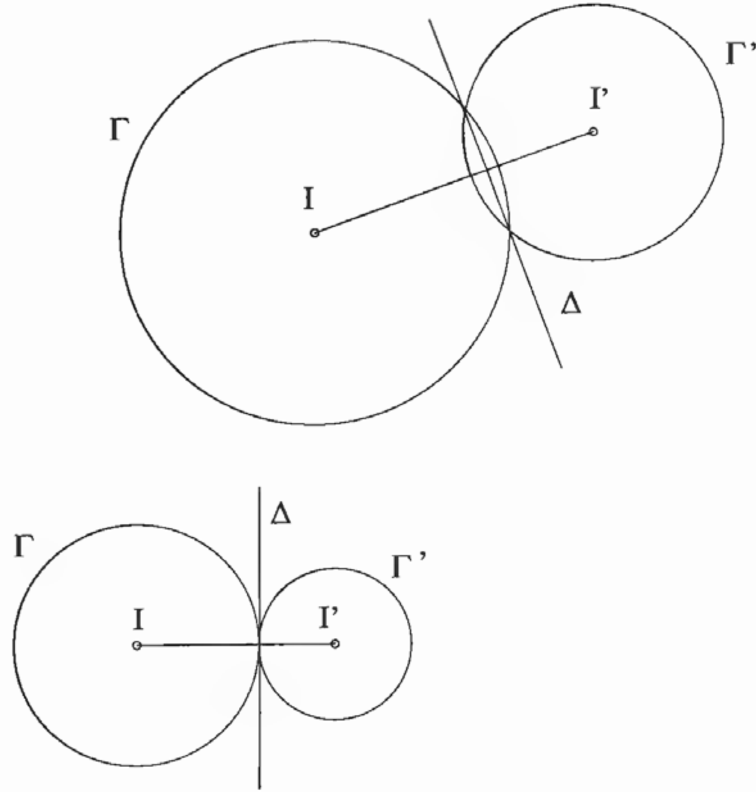



FIG. 6.11: *Axe radical de deux cercles sécants ou tangents.*

**Démonstration.**

Dans le repère orthonormé $(O, (\mathbf{e_1}, \mathbf{e_2}))$ les cercles $\Gamma$ et $\Gamma'$ ont pour équations respectives $x^2 + y^2 - 2ax - 2by + c = 0, \quad x^2 + y^2 - 2a'x - 2b'y + c' = 0$, et l'on a pour un point $M$ de coordonnées $(x, y)$ : $\mathcal{P}_\Gamma(M) = x^2 + y^2 - 2ax - 2by + c$, $\mathcal{P}_{\Gamma'}(M) = x^2 + y^2 - 2a'x - 2b'y + c'$. Donc

$$\Delta = \{M;\ 2(a' - a)x + 2(b' - b)y + c - c' = 0\}.$$

C'est une droite perpendiculaire au vecteur

$$\mathbf{II'} = \begin{pmatrix} a' - a \\ b' - b \end{pmatrix},$$

$I$ et $I'$ de coordonnées respectives $(a, b)$ et $(a', b')$ désignant les centres des cercles $\Gamma$ et $\Gamma'$. □

**Centre radical de 3 cercles, application à la construction de l'axe radical de 2 cercles d'intersection vide.**

**Proposition 6. 4.** *Soient $\Gamma, \Gamma', \Gamma"$ 3 cercles dont les centres respectifs $I, I', I"$ sont non alignés (en particulier deux-à-deux distincts). Soient $\Delta$ l'axe radical de $\Gamma$ et $\Gamma'$, $\Delta'$ l'axe radical de $\Gamma'$ et $\Gamma"$, $\Delta"$ l'axe radical de $\Gamma"$ et $\Gamma$. Alors les 3 axes $\Delta, \Delta', \Delta"$ sont concourants.*

**Démonstration.**

Par définition, $\Delta$ est l'ensemble des points $M$ tels que $\mathcal{P}_\Gamma(M) = \mathcal{P}_{\Gamma'}(M)$. De même, $\Delta'$ est l'ensemble des points $M$ tels que $\mathcal{P}_{\Gamma'}(M) = \mathcal{P}_{\Gamma"}(M)$ et $\Delta"$ est l'ensemble des points $M$ tels que $\mathcal{P}_{\Gamma"}(M) = \mathcal{P}_\Gamma(M)$. Comme les 3 centres ne sont pas

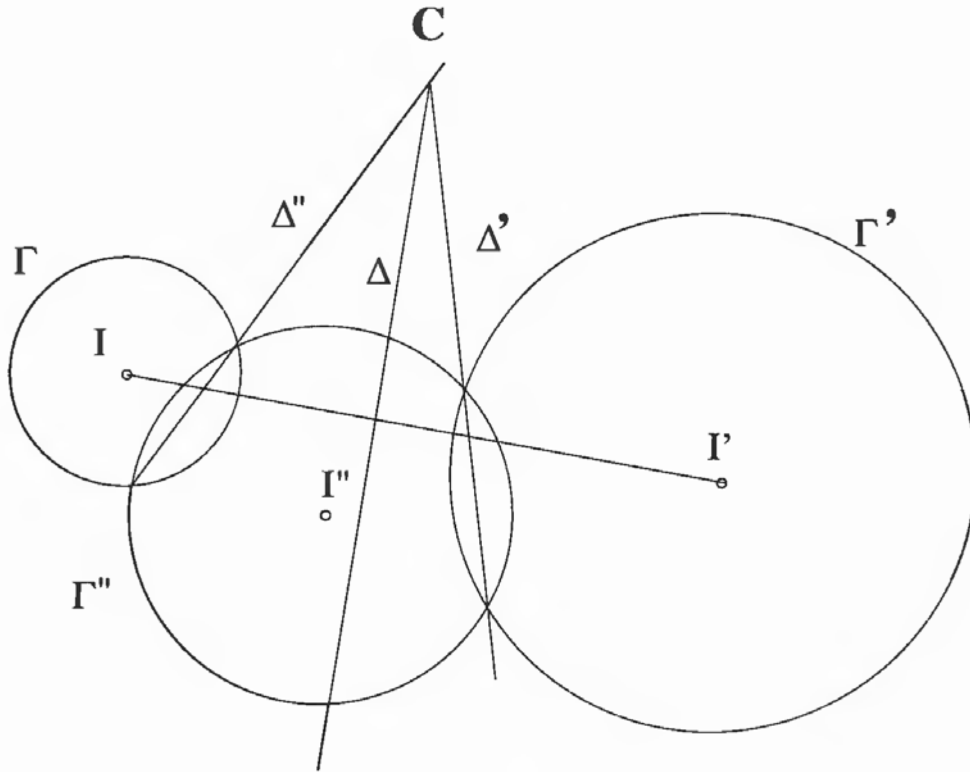



FIG. 6.12: *Centre radical de 3 cercles.*

alignés, $\Delta$ et $\Delta'$ ne sont pas parallèles. Il est alors immédiat que leur point d'intersection $M$ vérifie $\mathcal{P}_\Gamma(M) = \mathcal{P}_{\Gamma''}(M)$. Il est donc sur $\Delta''$. □

**Définition 6. 9.** *Dans les conditions de la proposition précédente, le point de concours des 3 axes radicaux s'appelle le centre radical des 3 cercles* $\Gamma, \Gamma', \Gamma''$.

**Application.** Soient $\Gamma$ et $\Gamma'$ deux cercles de centres respectifs $I, I'$ et dont l'intersection est vide. On prend un cercle $\Gamma''$ dont le centre $I''$ n'est pas aligné avec $I$ et $I'$, et tel que $\Gamma \cap \Gamma'' \neq \varnothing$ et $\Gamma' \cap \Gamma'' \neq \varnothing$ . Le point d'intersection $C$ des axes radicaux $\Delta''$ de $\Gamma$ et $\Gamma''$, et $\Delta'$ de $\Gamma'$ et $\Gamma''$, est le centre radical de 3 cercles. C'est donc un point de l'axe radical $\Delta$ de $\Gamma$ et $\Gamma'$. Par conséquent $\Delta$ est la perpendiculaire à la droite des centres $II'$ passant par $C$ (voir Fig. 6.12).

## 6.2.4 Faisceaux de cercles.

### Définition d'un faisceau de cercles.

Soient $\Gamma$ et $\Gamma'$ deux cercles, de centres $I$ et $I'$ distincts. Leurs équations respectives sont

$$\mathcal{P}_\Gamma(M) = 0, \quad \mathcal{P}_{\Gamma'}(M) = 0.$$

On considère l'ensemble des points $M$ tels que

$$\lambda\, \mathcal{P}_\Gamma(M) + \lambda'\, \mathcal{P}_{\Gamma'}(M) = 0 \qquad (6.34)$$

où $\lambda, \lambda' \in \mathbb{R}$ sont des paramètres non tous deux nuls.

Si $\lambda + \lambda' = 0$, (6.34) est l'équation de l'axe radical $\Delta$ de $\Gamma$ et $\Gamma'$.

Si $\lambda + \lambda' \neq 0$, en divisant (6.34) par $\lambda + \lambda'$, on peut supposer $\lambda + \lambda' = 1$, *i.e.* $\lambda' = 1 - \lambda$. L'ensemble considéré a donc pour équation

$$\lambda \mathcal{P}_\Gamma(M) + (1-\lambda)\mathcal{P}_{\Gamma'}(M) = 0 \tag{6.35}$$

avec $\lambda$ paramètre réel. Introduisons un repère affine orthonormé $(O, (\mathbf{i}, \mathbf{j}))$ tel que le point $O$ soit le point d'intersection de la ligne des centres $II'$ avec l'axe radical $\Delta$ de $\Gamma$ et $\Gamma'$, que l'axe des $x$ soit la ligne des centres et l'axe des $y$ l'axe radical. Si $M$ est un point de coordonnées $(x, y)$ dans ce repère,

$$\mathcal{P}_\Gamma(M) = x^2 + y^2 - 2ax + \gamma$$

$$\mathcal{P}_{\Gamma'}(M) = x^2 + y^2 - 2a'x + \gamma'$$

où $(a, 0)$ et $(a', 0)$ sont les coordonnées de $I$ et $I'$ respectivement. Alors $\mathcal{P}_\Gamma(O) = \gamma$ et $\mathcal{P}_{\Gamma'}(O) = \gamma'$. Or $O \in \Delta$, donc $\mathcal{P}_\Gamma(O) = \mathcal{P}_{\Gamma'}(O)$, *i.e.* $\gamma = \gamma'$. L'équation (6.35) s'écrit alors :

$$x^2 + y^2 - 2(\lambda a + (1-\lambda)a')x + \gamma = 0. \tag{6.36}$$

L'ensemble défini par cette équation est un cercle $\Gamma_\lambda$ si

$$(\lambda a + (1-\lambda)a')^2 \geqslant \gamma \tag{6.37}$$

et est l'ensemble vide sinon.

**Définition 6. 10.** *Soient $\Gamma$ et $\Gamma'$ deux cercles non concentriques, d'équations respectives*

$$\mathcal{P}_\Gamma(M) = 0, \quad \mathcal{P}_{\Gamma'}(M) = 0.$$

*On appelle faisceau de cercles défini par $\Gamma$ et $\Gamma'$ l'ensemble de tous les cercles qui ont une équation de la forme (6.35) pour un certain $\lambda \in \mathbb{R}$.*

Le faisceau de cercles défini par $\Gamma$ et $\Gamma'$ est donc l'ensemble des cercles $\Gamma_\lambda$ dont l'équation dans le repère $(O, (\mathbf{i}, \mathbf{j}))$ est (6.36), avec $\lambda$ astreint à vérifier la condition (6.37).

Le centre du cercle $\Gamma_\lambda$ est le point $I_\lambda$ de coordonnées $(\lambda a + (1-\lambda)a', 0)$. Le point $I_\lambda$ est le barycentre des points $I$ et $I'$ affectés des coefficients respectifs $\lambda$ et $1 - \lambda$. Si $\Gamma_{\lambda_1}$ et $\Gamma_{\lambda_2}$ ($\lambda_1 \neq \lambda_2$) sont deux cercles distincts du faisceau, $I_{\lambda_1} \neq I_{\lambda_2}$ puisque l'équation $\lambda_1 a + (1-\lambda_1)a' = \lambda_2 a + (1-\lambda_2)a'$, qui s'écrit $(\lambda_1 - \lambda_2)a = (\lambda_1 - \lambda_2)a'$, est impossible car $a \neq a'$. Si la condition (6.37) est vérifiée pour tout $\lambda \in \mathbb{R}$, quand $\lambda$ varie, le barycentre $I_\lambda$ peut être un point quelconque de la droite $II'$. Le rayon de $\Gamma_\lambda$ est

$$R_\lambda = \sqrt{(\lambda a + (1-\lambda)a')^2 - \gamma}. \tag{6.38}$$

**Proposition 6. 5.** *Considérons le faisceau de cercles défini par deux cercles non concentriques $\Gamma$ et $\Gamma'$ d'axe radical $\Delta$.*

*(i) Un cercle $\mathcal{C}$ appartient au faisceau si et seulement si il n'est pas concentrique avec $\Gamma$ et $\Delta$ est l'axe radical de $\Gamma$ et $\mathcal{C}$.*

*(ii) Tout couple de cercles distincts du faisceau a pour axe radical $\Delta$.*

*(iii) Si $\Gamma_{\lambda_1}$ et $\Gamma_{\lambda_2}$ ($\lambda_1 \neq \lambda_2$) sont deux cercles distincts du faisceau, le faisceau défini par $\Gamma_{\lambda_1}$ et $\Gamma_{\lambda_2}$ est le même que le faisceau défini par $\Gamma$ et $\Gamma'$.*

**Démonstration.**

(i) On a vu que si $\Gamma_{\lambda_1}$ et $\Gamma_{\lambda_2}$ $(\lambda_1 \neq \lambda_2)$ sont deux cercles distincts du faisceau, $I_{\lambda_1} \neq I_{\lambda_2}$. En prenant $\lambda_2 = 1$, *i.e.* $\Gamma_{\lambda_2} = \Gamma$, on obtient que tout cercle du faisceau distinct de $\Gamma$ est non concentrique avec $\Gamma$. Maintenant, soit $\mathcal{C}$ un cercle non concentrique avec $\Gamma$. L'axe radical de $\Gamma$ et $\Gamma'$ a pour équation

$$\mathcal{P}_\Gamma(M) - \mathcal{P}_{\Gamma'}(M) = 0.$$

L'axe radical de $\mathcal{C}$ et $\Gamma$ a pour équation

$$\mathcal{P}_{\mathcal{C}}(M) - \mathcal{P}_\Gamma(M) = 0.$$

Or deux droites coïncident si et seulement si leurs équations sont proportionnelles. Donc les axes radicaux coïncident si et seulement si il existe $\nu \in \mathbb{R}^*$ tel que

$$\mathcal{P}_{\mathcal{C}}(M) - \mathcal{P}_\Gamma(M) = \nu\,(\mathcal{P}_\Gamma(M) - \mathcal{P}_{\Gamma'}(M)) \ \ \forall\, M,$$

*i.e.*

$$\mathcal{P}_{\mathcal{C}}(M) = (1+\nu)\mathcal{P}_\Gamma(M) - \nu\,\mathcal{P}_{\Gamma'}(M) \ \ \forall\, M.$$

Cela s'écrit avec $\lambda = 1 + \nu$ :

$$\mathcal{P}_{\mathcal{C}}(M) = \lambda\,\mathcal{P}_\Gamma(M) + (1-\lambda)\mathcal{P}_{\Gamma'}(M) \ \ \forall\, M.$$

Donc les axes radicaux coïncident si et seulement si il existe $\lambda \neq 1$ tel que $\mathcal{C} = \Gamma_\lambda$.

(ii) Soient $\Gamma_{\lambda_1}, \Gamma_{\lambda_2}$ deux cercles du faisceau. Ils ne sont pas concentriques. D'après (i), $\Delta$ est l'axe radical de $\Gamma_{\lambda_1}$ et $\Gamma$. C'est aussi celui de $\Gamma_{\lambda_2}$ et $\Gamma$. Donc les points de $\Delta$ ont même puissance par rapport à $\Gamma_{\lambda_1}$ et $\Gamma_{\lambda_2}$ ce qui implique que l'axe radical de $\Gamma_{\lambda_1}$ et $\Gamma_{\lambda_2}$ est $\Delta$.

(iii) Considérons un cercle du faisceau défini par $\Gamma_{\lambda_1}$ et $\Gamma_{\lambda_2}$. Son équation est $\lambda\,\mathcal{P}_{\Gamma_{\lambda_1}}(M) + (1-\lambda)\,\mathcal{P}_{\Gamma_{\lambda_2}}(M) = 0$, pour un certain $\lambda \in \mathbb{R}$, *i.e.*

$$(\lambda\lambda_1 + (1-\lambda)\lambda_2)\,\mathcal{P}_\Gamma(M) + (\lambda(1-\lambda_1) + (1-\lambda)(1-\lambda_2))\mathcal{P}_{\Gamma'}(M) = 0.$$

Comme $\lambda\lambda_1 + (1-\lambda)\lambda_2 + \lambda(1-\lambda_1) + (1-\lambda)(1-\lambda_2) = 1$, le cercle appartient au faisceau défini par $\Gamma$ et $\Gamma'$. Réciproquement, on a par définition de la puissance $\mathcal{P}_{\Gamma_{\lambda_1}} = \lambda_1\,\mathcal{P}_\Gamma + (1-\lambda_1)\,\mathcal{P}_{\Gamma'}$ et $\mathcal{P}_{\Gamma_{\lambda_2}} = \lambda_2\,\mathcal{P}_\Gamma + (1-\lambda_2)\,\mathcal{P}_{\Gamma'}$. On en déduit $(1-\lambda_2)\,\mathcal{P}_{\Gamma_{\lambda_1}} - (1-\lambda_1)\,\mathcal{P}_{\Gamma_{\lambda_2}} = (\lambda_1 - \lambda_2)\,\mathcal{P}_\Gamma$ et $\lambda_2\,\mathcal{P}_{\Gamma_{\lambda_1}} - \lambda_1\,\mathcal{P}_{\Gamma_{\lambda_2}} = (\lambda_2 - \lambda_1)\,\mathcal{P}_{\Gamma'}$, donc, puisque $\lambda_1 - \lambda_2 \neq 0$, $\Gamma$ et $\Gamma'$ appartiennent au faisceau défini par $\Gamma_{\lambda_1}$ et $\Gamma_{\lambda_2}$. Il en résulte alors comme plus haut que tout cercle du faisceau défini par $\Gamma$ et $\Gamma'$ appartient au faisceau défini par $\Gamma_{\lambda_1}$ et $\Gamma_{\lambda_2}$. Les deux faisceaux sont donc égaux. □

**Définition 6. 11.** *Considérons le faisceau de cercles défini par deux cercles non concentriques $\Gamma$ et $\Gamma'$ d'axe radical $\Delta$. On appelle $\Delta$ l'axe radical du faisceau.*

**Classification des faisceaux de cercles.**

Considérons le faisceau de cercles défini par les deux cercles non concentriques $\Gamma$ et $\Gamma'$. Il y a 3 cas à distinguer.

► $\gamma < 0$. Dans ce cas la condition (6.37) est réalisée quel que soit $\lambda \in \mathbb{R}$. De plus $\mathcal{P}_\Gamma(O) < 0$ et $\mathcal{P}_{\Gamma'}(O) < 0$, donc le point $O$ est intérieur aux deux cercles $\Gamma$ et $\Gamma'$. L'axe radical de $\Gamma$ et $\Gamma'$ rencontre donc $\Gamma$ en 2 points $A$ et $B$ qui sont alors nécessairement les points d'intersection de $\Gamma$ et $\Gamma'$. Pour tout $\lambda \in \mathbb{R}$, le cercle

$\Gamma_\lambda$ passe par $A$ et par $B$ puisque $\mathcal{P}_\Gamma(A) = \mathcal{P}_{\Gamma'}(A) = 0$ et $\mathcal{P}_\Gamma(B) = \mathcal{P}_{\Gamma'}(B) = 0$. Réciproquement, si un cercle $\mathcal{C}$ passe par $A$ et par $B$, les points $A$ et $B$ sont communs au 3 cercles $\Gamma, \Gamma'$ et $\mathcal{C}$, donc la droite $AB$ est l'axe radical de deux quelconques d'entre eux. En particulier, le centre de $\mathcal{C}$ est situé sur la droite $II'$. Le centre de $\mathcal{C}$ est donc $I_\lambda$ pour un certain $\lambda$. Les deux cercles $\mathcal{C}$ et $\Gamma_\lambda$ ont le même centre et passent tous les deux par les points $A$ et $B$, donc il coïncident.

Ainsi la famille des cercles $\Gamma_\lambda$, $\lambda \in \mathbb{R}$, est celle de tous les cercles passant par les deux points $A$ et $B$. L'axe radical du faisceau est la droite $AB$. On dit que le faisceau est un *faisceau de cercles sécants.*

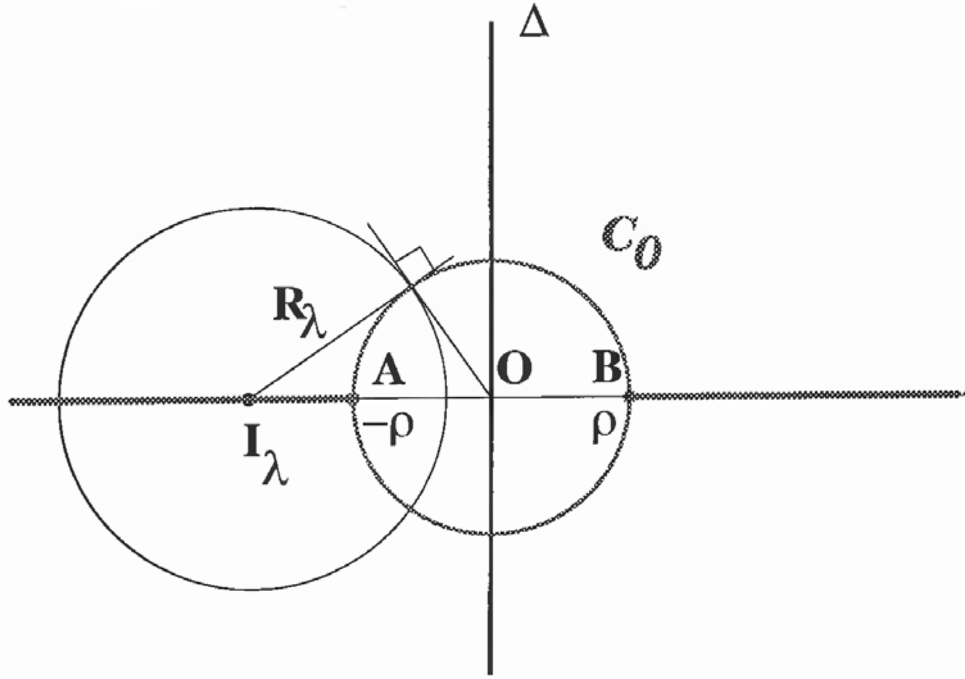



FIG. 6.13: *Faisceau à points limites $A$ et $B$, avec $\rho = \sqrt{\gamma}$.*

► $\gamma = 0$. La condition (6.37) est encore réalisée quel que soit $\lambda \in \mathbb{R}$. De plus $\Gamma$ et $\Gamma'$ passent par $O$, et comme $O$ est sur la droite $II'$, ils sont tangents en $O$ à $\Delta$. $\Gamma_\lambda$ passe aussi par $O$, et, étant centré sur l'axe de $x$, il est aussi tangent en $O$ à $\Delta$. Réciproquement, si $\mathcal{C}$ est un tel cercle, son centre est $I_\lambda$ pour un certain $\lambda$. Les deux cercles $\mathcal{C}$ et $\Gamma_\lambda$ ont le même centre et passent tous les deux par l'origine, donc il coïncident.

Ainsi la famille des cercles $\Gamma_\lambda$, $\lambda \in \mathbb{R}$, est celle de tous les cercles passant par l'origine $O$ et tangents en $O$ à $\Delta$. C'est aussi la famille de tous les cercles tangents en $O$ et dont deux quelconques d'entre eux ont pour axe radical $\Delta$. On dit que le faisceau est un *faisceau de cercles tangents dont l'axe radical est* $\Delta$.

► $\gamma > 0$. Dans ce cas le point $O$ est extérieur au deux cercles $\Gamma$ et $\Gamma'$. Cela implique que $\Gamma \cap \Gamma' = \varnothing$. La condition (6.37) n'est pas toujours réalisée. Elle s'écrit

$$|\lambda a + (1 - \lambda)a'| \geqslant \sqrt{\gamma}. \tag{6.39}$$

L'abscisse de $I_\lambda$ est donc dans l'intervalle $]-\infty, -\sqrt{\gamma}]$ ou dans l'intervalle $[\sqrt{\gamma}, +\infty[$, et peut être un point quelconque de ces intervalles. Le rayon du cercle $\Gamma_\lambda$ est donné par la formule (6.38) *i.e.*

$$R_\lambda^2 + \gamma = (\lambda a + (1 - \lambda)a')^2 = \|\mathbf{OI}_\lambda\|^2. \tag{6.40}$$

En particulier, on voit que lorsque $I_\lambda$ est l'un des points $A$ ou $B$ de coordonnées respectives $(-\sqrt{\gamma}, 0)$ et $(\sqrt{\gamma}, 0)$, le rayon de $\Gamma_\lambda$ est *nul*. Les points $A$ et $B$ sont des points *limites*, et on dit que le faisceau est un *faisceau à points limites*. D'après la Prop. 6.5, le faisceau à points limites $A, B$ est aussi le faisceau défini par les deux cercle-points (de rayon nul) $\{A\}$ et $\{B\}$.

Soit $\mathcal{C}_0$ le cercle de centre $O$ et de rayon $\sqrt{\gamma}$. Un cercle $\mathcal{C}$ appartient au faisceau si et seulement si son centre $I$ est sur l'axe des $x$ et est lié à son rayon par l'équation (6.40). Cela signifie que le cercle est orthogonal au cercle $\mathcal{C}_0$. Le faisceau est donc l'ensemble des cercles orthogonaux au cercle $\mathcal{C}_0$ et dont le centre est sur l'axe des $x$ (voir Fig. 6.13).

# 6.3 Sections planes d'un cône de révolution.

Le traitement analytique de cette section est inspiré de [1], p.197. Pour une démonstration purement géométrique utilisant les sphères inscrites tangentes au plan sécant (démonstration de Dandelin), voir [12], p. 184 - 187.

Soit $(O, (\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}))$ un repère orthonormé de $\mathbb{R}^3$ et $\mathcal{C}$ un cône de révolution de sommet l'origine $O$, et dont l'axe est le vecteur $\mathbf{e_3}$, ayant pour ouverture $\alpha$ $(0 < \alpha < \frac{\pi}{2})$. L'équation du cône $\mathcal{C}$ est

$$(\mathbf{OM}|\mathbf{e_3})^2 = \|\mathbf{OM}\|^2 \cos^2\alpha \tag{6.41}$$

ou encore en notant $(x, y, z)$ les coordonnées de $M$

$$z^2 \tan^2\alpha = x^2 + y^2.$$

Le cône $\mathcal{C}$ a deux nappes dont les équations sont

$$(\mathbf{OM}|\mathbf{e_3}) = \epsilon \|\mathbf{OM}\| \cos\alpha \tag{6.42}$$

avec $\epsilon = 1$ pour la nappe supérieure (correspondant à $z \geqslant 0$) et $\epsilon = -1$ pour la nappe inférieure (correspondant à $z \leqslant 0$).

Soit $\Pi$ un plan ne passant pas par $O$ et $\mathbf{f_3}$ un vecteur unitaire normal à $\Pi$, orienté de sorte que l'angle non orienté $\beta$ des vecteurs $\mathbf{e_3}$ et $\mathbf{f_3}$ défini par $(\mathbf{e_3}|\mathbf{f_3}) = \cos\beta$, $0 \leqslant \beta \leqslant \pi$, vérifie $0 \leqslant \beta \leqslant \frac{\pi}{2}$ (voir Fig. 6.14). Par rotation des axes $x, y$ autour de l'axe des $z$, on peut supposer que le plan $(O, \mathbf{e_3}, \mathbf{f_3})$ est le plan $x, z$.

Soit $\Delta$ la droite intersection de $\Pi$ avec le plan $(O, \mathbf{e_3}, \mathbf{f_3})$, $\mathbf{f_1}$ un vecteur unitaire de $\Delta$, et $\mathbf{f_2}$ tel que $(\mathbf{f_1}, \mathbf{f_2}, \mathbf{f_3})$ soit une base orthonormée.

Soit $O'$ l'un des points d'intersection de $\Delta$ avec le cône. Nous utiliserons le repère orthonormé $(O', (\mathbf{f_1}, \mathbf{f_2}, \mathbf{f_3}))$ (voir Fig. 6.15 dans le plan $(O, \mathbf{e_3}, \mathbf{f_3})$). On notera $(\xi, \eta, \zeta)$ les coordonnées par rapport à ce repère. Dans ce repère, on a $\mathbf{e_3} = \begin{pmatrix} \varepsilon \sin\beta \\ 0 \\ \cos\beta \end{pmatrix}$, avec $\varepsilon = \pm 1$. $\varepsilon$ dépend du choix de l'orientation du vecteur $\mathbf{f_1}$. On peut supposer ce vecteur choisi de sorte que $\varepsilon = 1$.

Ainsi $\mathbf{e_3} = \begin{pmatrix} \sin\beta \\ 0 \\ \cos\beta \end{pmatrix}$, $\mathbf{O'O} = \begin{pmatrix} a \\ 0 \\ c \end{pmatrix}$ $(a, c \in \mathbb{R})$, $\mathbf{O'M} = \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix}$ et $\mathbf{OM} = \begin{pmatrix} \xi - a \\ \eta \\ \zeta - c \end{pmatrix}$. L'équation du cône (6.41) s'écrit donc :

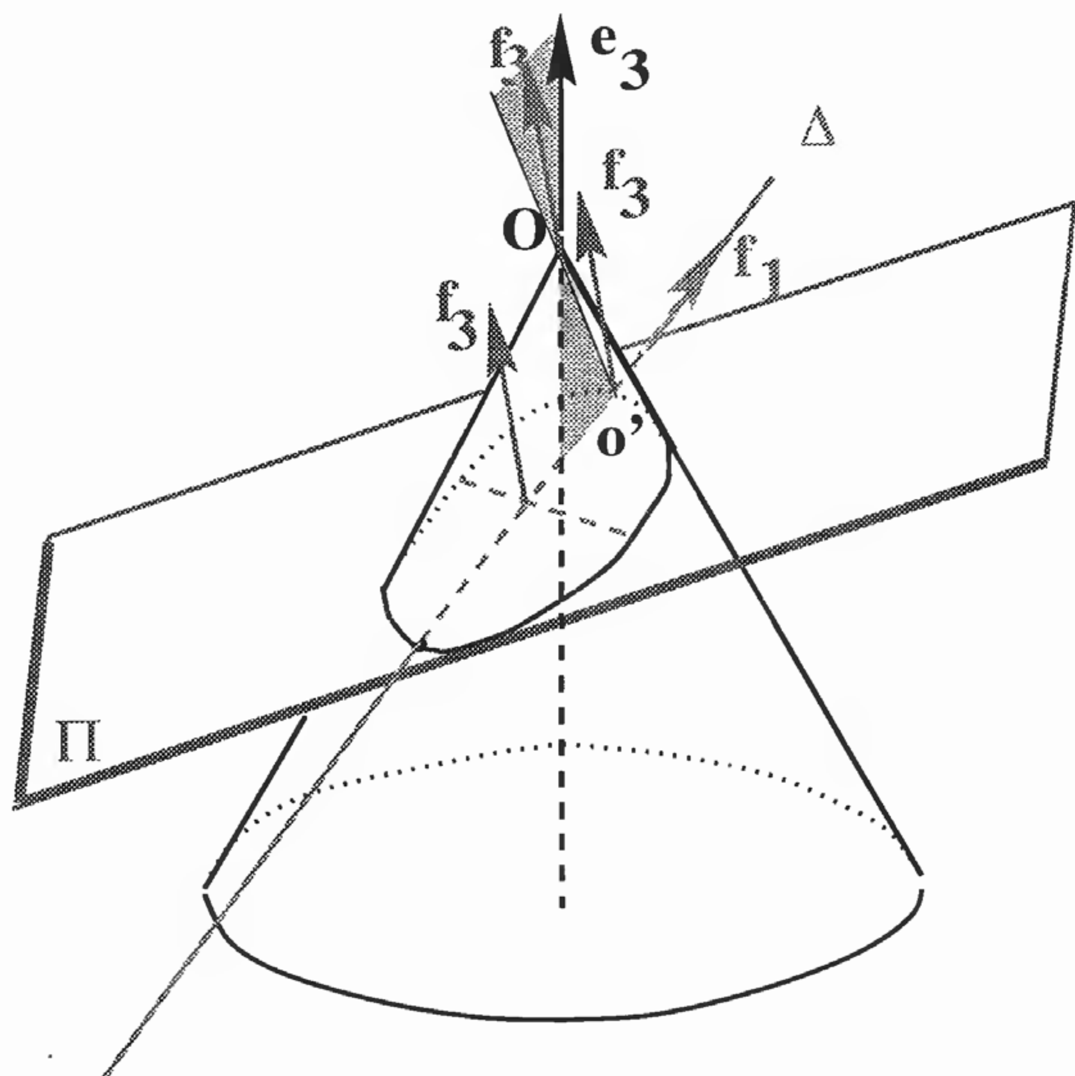



FIG. 6.14: *Section d'un cône de révolution.*

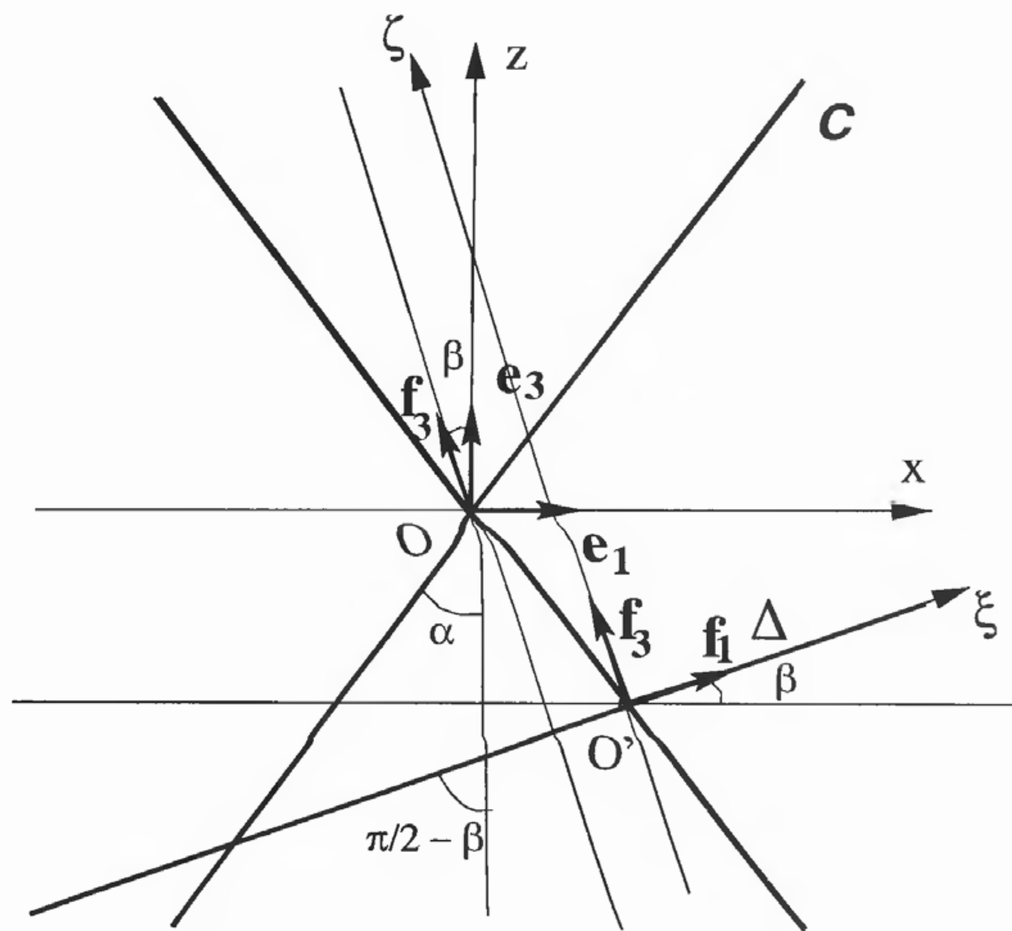



FIG. 6.15: *Section d'un cône de révolution. Figure dans le plan* $(O, \mathbf{e_3}, \mathbf{f_3})$.

$$((\xi - a)\sin\beta + (\zeta - c)\cos\beta)^2 = ((\xi - a)^2 + \eta^2 + (\zeta - c)^2)\cos^2\alpha.$$

L'équation de l'intersection du cône $\mathcal{C}$ avec le plan $\Pi$ est obtenue en faisant $\zeta = 0$ dans l'équation ci-dessus. C'est donc :

$$((\xi - a)\sin\beta - c\cos\beta)^2 = ((\xi - a)^2 + \eta^2 + c^2)\cos^2\alpha. \tag{6.43}$$

Cette équation s'écrit

$$\xi^2\sin^2\beta - 2\xi\sin\beta(a\sin\beta + c\cos\beta) + (a\sin\beta + c\cos\beta)^2 = (\xi^2 - 2\xi a + \eta^2 + a^2 + c^2)\cos^2\alpha$$

ou encore

$$\xi^2(\cos^2\alpha - \sin^2\beta) + 2\xi((\mathbf{O'O}|\mathbf{e_3})\sin\beta - a\cos^2\alpha) + \eta^2\cos^2\alpha = 0 \tag{6.44}$$

puisque

$$(a\sin\beta + c\cos\beta)^2 - (a^2 + c^2)\cos^2\alpha = (\mathbf{O'O}|\mathbf{e_3})^2 - \|\mathbf{O'O}\|^2\cos^2\alpha = 0$$

car le point $O'$ est sur le cône donc vérifie l'équation (6.41).

Comme d'après (6.42)

$$(\mathbf{O'O}|\mathbf{e_3}) = -(\mathbf{OO'}|\mathbf{e_3}) = -\epsilon\|\mathbf{O'O}\|\cos\alpha$$

avec $\epsilon = \pm 1$ suivant la nappe de cône sur laquelle est le point $O'$, l'équation (6.44) s'écrit en posant $e = \frac{\sin\beta}{\cos\alpha}$ :

$$\xi^2(1 - e^2) - 2\xi(\epsilon\|\mathbf{O'O}\|e + a) + \eta^2 = 0.$$

Si l'on pose $p = \epsilon\|\mathbf{O'O}\|e + a$, on obtient finalement

$$\xi^2(1 - e^2) - 2\xi p + \eta^2 = 0. \tag{6.45}$$

• **1er cas :** $e = 1$ , *i.e.* $\frac{\pi}{2} - \beta = \alpha$. L'équation (6.45) est :

$$\eta^2 = 2p\xi.$$

C'est l'équation d'une *parabole*. La droite $\Delta$ est parallèle à l'une des deux génératrices du cône située dans le plan $(\xi, \zeta)$, et $O'$ est l'unique point d'intersection de $\Delta$ avec le cône. On a $p < 0$ si $O'$ est sur la nappe inférieure ($\epsilon = -1$) et $p > 0$ si $O'$ est sur la nappe supérieure ($\epsilon = +1$). La parabole a pour paramètre $|p|$, elle est dirigée vers les $\xi < 0$ si $O'$ est sur la nappe inférieure et vers les $\xi > 0$ si $O'$ est sur la nappe supérieure.

• **2ème cas :** $e \neq 1$, *i.e.* $\frac{\pi}{2} - \beta \neq \alpha$. L'équation (6.45) s'écrit dans ce cas :

$$\xi^2 - 2\xi\frac{p}{1 - e^2} + \frac{\eta^2}{1 - e^2} = 0$$

donc

$$(\xi - \frac{p}{1 - e^2})^2 + \frac{\eta^2}{1 - e^2} = \frac{p^2}{(1 - e^2)^2}$$

ou encore en posant $X = \xi - \frac{p}{1-e^2}$ et $Y = \eta$ :

$$X^2 + \frac{Y^2}{1 - e^2} = \frac{p^2}{(1 - e^2)^2}. \tag{6.46}$$

Si $e < 1$ , *i.e.* $\frac{\pi}{2} - \beta > \alpha$, l'équation (6.46) s'écrit :

$$\frac{X^2}{a^2} + \frac{Y^2}{b^2} = 1 \tag{6.47}$$

avec $a = \frac{|p|}{1-e^2}$ et $b = \frac{|p|}{\sqrt{1-e^2}} = a\sqrt{1-e^2}$. C'est l'équation d'une *ellipse* dont le grand axe est $a$ et le petit axe $b$.

Si $e > 1$ , *i.e.* $\frac{\pi}{2} - \beta < \alpha$, l'équation (6.46) s'écrit :

$$\frac{X^2}{a^2} - \frac{Y^2}{b^2} = 1 \tag{6.48}$$

avec $a = \frac{|p|}{e^2-1}$ et $b = \frac{|p|}{\sqrt{e^2-1}} = a\sqrt{e^2-1}$. C'est l'équation d'une *hyperbole* dont le grand axe est $a$ et le petit axe $b$.

## 6.4 Mouvement à accélération centrale en $\frac{1}{r^2}$.

### 6.4.1 Énergie.

**Repère galiléen.**

Dans l'espace affine euclidien $\mathbb{R}^3$, soit $\mathcal{R} = (O, (\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}))$ un repère orthonormé. Les fonctions introduites sont supposées de classe $C^2$.

Si $\mathbf{v}(t) = x(t)\,\mathbf{e_1} + y(t)\,\mathbf{e_2} + z(t)\,\mathbf{e_3}$ est un vecteur de l'espace vectoriel $\mathbb{R}^3$, nous utiliserons la notation $\left(\frac{d\mathbf{OM}}{dt}\right)_\mathcal{R}$ pour désigner le vecteur $\frac{dx}{dt}\,\mathbf{e_1} + \frac{dy}{dt}\,\mathbf{e_2} + \frac{dz}{dt}\,\mathbf{e_3}$ lorsqu'il faudra souligner que l'on dérive par rapport au temps $t$ les composantes dans $\mathcal{R}$ du vecteur $\mathbf{v}$.

On dit que $\mathcal{R}$ est un *repère galiléen* si le mouvement par rapport à $\mathcal{R}$ d'un point matériel quelconque $M(t)$ de masse $m$ soumis à chaque instant $t$ à une force $\mathbf{F}(t)$ vérifie le *principe fondamental de la dynamique du point matériel* :

$$\mathbf{F} = m\mathbf{\Gamma} \tag{6.49}$$

où $\mathbf{\Gamma} = \left(\frac{d^2\mathbf{OM}}{dt^2}\right)_\mathcal{R}$ désigne l'accélération de $M$ par rapport à $\mathcal{R}$.

Soit $\mathcal{R} = (O, (\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}))$ un repère galiléen et $\mathcal{R}_1 = (O_1(t), (\mathbf{f_1}(t), \mathbf{f_2}(t), \mathbf{f_3}(t)))$ un repère orthonormé mobile quelconque. Si $(\xi(t), \eta(t), \zeta(t))$ désignent les coordonnées dans $\mathcal{R}_1$ d'un point mobile $M(t)$, on a

$$\begin{aligned}\mathbf{O_1M} &= \xi\,\mathbf{f_1} + \eta\,\mathbf{f_2} + \zeta\,\mathbf{f_3}\\ \mathbf{OM} &= \mathbf{OO_1} + \xi\mathbf{f_1} + \eta\mathbf{f_2} + \zeta\mathbf{f_3}\end{aligned}$$

donc

$$\begin{aligned}\left(\frac{d\mathbf{OM}}{dt}\right)_\mathcal{R} = \left(\frac{d\mathbf{OO_1}}{dt}\right)_\mathcal{R} &+ \frac{d\xi}{dt}\,\mathbf{f_1} + \frac{d\eta}{dt}\,\mathbf{f_2} + \frac{d\zeta}{dt}\,\mathbf{f_3}\\ &+ \xi\left(\frac{d\mathbf{f_1}}{dt}\right)_\mathcal{R} + \eta\left(\frac{d\mathbf{f_2}}{dt}\right)_\mathcal{R} + \zeta\left(\frac{d\mathbf{f_3}}{dt}\right)_\mathcal{R}.\end{aligned} \tag{6.50}$$

Cela s'écrit

$$\mathbf{V}(M/\mathcal{R}) = \mathbf{V}(M/\mathcal{R}_1) + \mathbf{V_e}(M/\mathcal{R}) \tag{6.51}$$

où $\mathbf{V}(M/\mathcal{R}) = \left(\frac{d\mathbf{OM}}{dt}\right)_\mathcal{R}$ est la vitesse de $M$ par rapport à $\mathcal{R}$,

$$\mathbf{V}(M/\mathcal{R}_1) = \left(\frac{d\mathbf{O_1M}}{dt}\right)_{\mathcal{R}_1} = \frac{d\xi}{dt}\mathbf{f_1} + \frac{d\eta}{dt}\mathbf{f_2} + \frac{d\zeta}{dt}\mathbf{f_3}$$

est la vitesse de $M$ par rapport à $\mathcal{R}_1$, et

$$\mathbf{V}_e(M/\mathcal{R}) = \left(\frac{d\mathbf{OO_1}}{dt}\right)_{\mathcal{R}} + \xi\left(\frac{d\mathbf{f_1}}{dt}\right)_{\mathcal{R}} + \eta\left(\frac{d\mathbf{f_2}}{dt}\right)_{\mathcal{R}} + \zeta\left(\frac{d\mathbf{f_3}}{dt}\right)_{\mathcal{R}}$$

est la vitesse au temps $t$ du point *fixe dans* $\mathcal{R}_1$ qui coïncide avec $M(t)$ au temps $t$ (ce point ne coïncide en général plus avec avec $M(s)$ au temps $s > t$). On dit que $\mathbf{V}_e(M/\mathcal{R})$ est la *vitesse d'entraînement.* Le mouvement de $\mathcal{R}_1$ par rapport à $\mathcal{R}$ est le *mouvement d'entraînement.*

Si le mouvement d'entraînement est un mouvement de translation rectiligne uniforme, *i.e.* tous les point de $\mathcal{R}_1$ ont la même vitesse constante $\mathbf{k}$ par rapport à $\mathcal{R}$, on a pour tous points $A, B$ fixés de $\mathcal{R}_1$

$$\left(\frac{d\mathbf{AB}}{dt}\right)_{\mathcal{R}} = \left(\frac{d\mathbf{OB}}{dt}\right)_{\mathcal{R}} - \left(\frac{d\mathbf{OA}}{dt}\right)_{\mathcal{R}} = \mathbf{k} - \mathbf{k} = 0.$$

En particulier $(\frac{d\mathbf{f_1}}{dt})_{\mathcal{R}} = (\frac{d\mathbf{f_2}}{dt})_{\mathcal{R}} = (\frac{d\mathbf{f_3}}{dt})_{\mathcal{R}} = 0$ donc

$$\left(\frac{d\mathbf{OM}}{dt}\right)_{\mathcal{R}} = \mathbf{k} + \frac{d\xi}{dt}\mathbf{f_1} + \frac{d\eta}{dt}\mathbf{f_2} + \frac{d\zeta}{dt}\mathbf{f_3}$$

et

$$\left(\frac{d^2\mathbf{OM}}{dt^2}\right)_{\mathcal{R}} = \frac{d^2\xi}{dt^2}\mathbf{f_1} + \frac{d^2\eta}{dt^2}\mathbf{f_2} + \frac{d^2\zeta}{dt^2}\mathbf{f_3} = \left(\frac{d^2\mathbf{O_1M}}{dt^2}\right)_{\mathcal{R}_1}$$

*i.e.* l'accélération $\mathbf{\Gamma}(M/\mathcal{R})$ de $M$ par rapport à $\mathcal{R}$ est égale à l'accélération $\mathbf{\Gamma}(M/\mathcal{R}_1)$ de $M$ par rapport à $\mathcal{R}_1$. Il en résulte que le mouvement de $M$ par rapport à $\mathcal{R}_1$ vérifie (6.49), donc $\mathcal{R}_1$ est galiléen.

Réciproquement, supposons que $\mathcal{R}_1$ soit galiléen. D'après le principe fondamental (6.49) valable pour $\mathcal{R}_1$ puisque $\mathcal{R}_1$ est galiléen, un point $A$ fixé de $\mathcal{R}_1$ n'est soumis à aucune force car son accélération par rapport $\mathcal{R}_1$ est nulle. Comme $A$ n'est soumis à aucune force et que $\mathcal{R}$ est galiléen, l'accélération de $A$ par rapport $\mathcal{R}$ est aussi nulle d'après le principe fondamental (6.49) valable pour $\mathcal{R}$. La vitesse $\mathbf{V}(A/\mathcal{R})$ de $A$ par rapport à $\mathcal{R}$ est donc un vecteur constant $\mathbf{V}_A$ et par intégration il existe un vecteur $\mathbf{k}_A$ tel que

$$\mathbf{OA} = t\,\mathbf{V}_A + \mathbf{k}_A \quad \forall t\,.$$

Pour tous point $A, B$ de $\mathcal{R}_1$, on aura donc

$$\mathbf{AB} = \mathbf{OB} - \mathbf{OA} = t\,(\mathbf{V}_B - \mathbf{V}_A) + \mathbf{k}_B - \mathbf{k}_A \quad \forall t\,.$$

D'où

$$\|\mathbf{AB}\|^2 = t^2\|\mathbf{V}_B - \mathbf{V}_A\|^2 + 2t(\mathbf{V}_B - \mathbf{V}_A|\mathbf{k}_B - \mathbf{k}_A) + \|\mathbf{k}_B - \mathbf{k}_A\|^2 \quad \forall t\,.$$

Comme $\|\mathbf{AB}\|^2$ est une constante puisque $A$ et $B$ sont fixes dans $\mathcal{R}_1$, nécessairement $\mathbf{V}_B - \mathbf{V}_A = 0$. Tous les points de $\mathcal{R}_1$ ont donc la même vitesse constante par rapport à $\mathcal{R}$. Le mouvement de $\mathcal{R}_1$ par rapport à $\mathcal{R}$ est un mouvement de translation rectiligne uniforme.

On a donc obtenu : *si $\mathcal{R}$ est un repère galiléen, les autres repères galiléens sont exactement les repères dont le mouvement par rapport à $\mathcal{R}$ est un mouvement de translation rectiligne uniforme.* A noter que l'existence d'un repère galiléen est un *postulat* de la Mécanique Classique. Le terme *galiléen* est choisi pour rappeler qu'on se place dans le cadre de la Mécanique Classique, où il est supposé exister un temps universel pour tous les observateurs, et non dans le cadre de la Mécanique Relativiste.

Décider qu'un repère donné est ou non galiléen est une question d'approximation.

**Exemples usuels de repères galiléens approchés.**

Le repère ayant pour origine le centre de gravité du système solaire et pour axes trois directions stellaires peut être considéré comme galiléen. C'est le *repère de Copernic.*

Le repère ayant pour origine le Soleil (assimilé à son centre de gravité) et pour axes trois directions stellaires est une bonne approximation de repère galiléen pour l'étude des planètes. C'est le *repère de Kepler.*

**Champ central.**

Dans l'espace affine euclidien $\mathbb{R}^3$, on considère le mouvement par rapport à un repère orthonormé $\mathcal{R} = (O, (\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}))$ d'un point $M(t)$ de masse $m$ sur lequel s'exerce un champ de forces central de pôle l'origine $O$, et en $\frac{1}{r^2}$. La force qui s'exerce sur le point $M$ est

$$\mathbf{F}_M = -\frac{k}{r^2}\,\mathbf{u} \tag{6.52}$$

où $\mathbf{u} = \frac{\mathbf{OM}}{\|\mathbf{OM}\|}, r = \|\mathbf{OM}\|$, et $k$ est une constante $\neq 0$. Si $k > 0$, il y a *attraction*; si $k < 0$, il y a *répulsion.*

Dans toute la suite de cette section, nous ferons l'hypothèse suivante.

$$\textit{Le repère } \mathcal{R} = (O, (\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3})) \textit{ est galiléen.} \tag{6.53}$$

**Exemples.**

1) Une masse $\mu$ placée en $O$ engendre dans $\mathbb{R}^3 \setminus \{O\}$ un champ de gravitation dont la valeur au point $M$ est

$$\mathbf{G}_M = -\frac{\mathcal{G}\mu}{r^2}\,\mathbf{u} \tag{6.54}$$

où $\mathcal{G}$ est la constante de gravitation universelle ($\mathcal{G} \approx 6.67 \cdot 10^{-11}$ dans le système international SI). Le champ de forces qui s'exerce sur le point $M$ de masse $m$ est alors :

$$\mathbf{F}_M = -\frac{\mathcal{G}\mu\, m}{r^2}\,\mathbf{u}. \tag{6.55}$$

Il y a attraction.

2) Une charge électrique $Q$ placée en $O$ engendre dans $\mathbb{R}^3 \setminus \{O\}$ un champ électrique (champ de Coulomb) dont la valeur au point $M$ est

$$\mathbf{E}_M = -\frac{Q}{4\pi\epsilon_0\, r^2}\,\mathbf{u} \tag{6.56}$$

où la constante $4\pi\epsilon_0$ est appelée *la permittivité du vide* ($4\pi\epsilon_0 \approx 1.113 \cdot 10^{-10}$ dans le système SI, soit $\frac{1}{4\pi\epsilon_0} \approx 8.9847 \cdot 10^9$ SI). Le champ de forces qui s'exerce sur le point $M$ ne dépend pas de la masse $m$ mais de la charge $q$ de $M$ et est alors :

$$\mathbf{F}_M = -\frac{Q\,q}{4\pi\epsilon_0\, r^2}\,\mathbf{u}. \tag{6.57}$$

Il y a attraction ou répulsion suivant que les charges $Q, q$ sont de signes contraires ou non.

**Remarque: problème à deux corps.**

Dans les deux exemples physiques précédents, l'hypothèse (6.53) n'est vérifiée qu'approximativement. On néglige en effet l'action de la masse (*resp :* charge) mobile $M(t)$ sur la masse (*resp :* charge) $O$. La loi de la gravitation universelle (*resp :* la loi de Coulomb) dit qu'il y a interaction. Le problème réel est un *problème à deux corps.* Nous allons examiner le problème à deux corps dans le cas de la gravitation, le cas du champ de Coulomb étant analogue.

On introduit un repère galiléen $\mathcal{R}_0 = (\Omega, (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3))$ auquel seront rapportés les mouvements de $O$ et $M$. Les deux masses ponctuelles $\mu$ placée en $O$ et $m$ placée en $M$ exercent l'une sur l'autre deux forces opposées. La force exercée par $O$ sur $M$ est $\mathbf{F}_M$ donnée par (6.55) et la force exercée par $M$ sur $O$ est $\mathbf{F}_O = -\mathbf{F}_M$.

► Mouvement par rapport au centre de gravité.

Soit $G$ le barycentre des points $O$ et $M$ affectés de leur masse respective. Par définition,

$$\mathbf{\Omega G} = \frac{1}{\mu + m}\,(\mu\,\mathbf{\Omega O} + m\,\mathbf{\Omega M}).$$

En dérivant 2 fois par rapport au temps $t$, on obtient

$$\mathbf{\Gamma}(G/\mathcal{R}_0) = \frac{1}{\mu + m}\,(\mu\,\mathbf{\Gamma}(O/\mathcal{R}_0) + m\,\mathbf{\Gamma}(M/\mathcal{R}_0)).$$

Or d'après le principe fondamental (6.49) valable pour $\mathcal{R}_0$,

$$\begin{aligned} \mu\,\mathbf{\Gamma}(O/\mathcal{R}_0) &= \mathbf{F}_O \\ m\,\mathbf{\Gamma}(M/\mathcal{R}_0) &= \mathbf{F}_M \end{aligned}$$

donc

$$\mathbf{\Gamma}(G/\mathcal{R}_0) = 0.$$

Il en résulte immédiatement que le mouvement du repère $\mathcal{R}_G = (G, (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3))$ par rapport à $\mathcal{R}_0$ est un mouvement de translation rectiligne uniforme, et donc que le repère $\mathcal{R}_G$ est galiléen. On a

$$\mathbf{GM} = \mathbf{\Omega M} - \mathbf{\Omega G} = \frac{\mu}{\mu + m}\,(\mathbf{\Omega M} - \mathbf{\Omega O}) = \frac{\mu}{\mu + m}\,\mathbf{OM}.$$

D'où

$$\|\mathbf{OM}\| = (1 + \frac{m}{\mu})\|\mathbf{GM}\|\,,$$

*i.e.*

$$r = (1 + \frac{m}{\mu})\,r_G,$$

avec $r = \|\mathbf{OM}\|$ et $r_G = \|\mathbf{GM}\|$. Donc

$$\mathbf{F}_M = -\frac{\mathcal{G}m\mu}{r^2}\,\mathbf{u} = -\frac{\mathcal{G}m\mu}{(1 + \frac{m}{\mu})^2\,r_G^2}\,\mathbf{u}.$$

Le repère $\mathcal{R}_G$ étant galiléen, on a alors

$$m\left(\frac{d^2\mathbf{GM}}{dt^2}\right)_{\mathcal{R}_G} = m\,\mathbf{\Gamma}(M/\mathcal{R}_G) = -\frac{\mathcal{G}m\mu}{(1 + \frac{m}{\mu})^2\,r_G^2}\,\mathbf{u}.$$

A noter que $\mathbf{u} = \frac{\mathbf{OM}}{\|\mathbf{OM}\|} = \frac{\mathbf{GM}}{\|\mathbf{GM}\|}$. L'équation du mouvement de $M$ dans le repère galiléen $\mathcal{R}_G$ est donc celle d'un mouvement sous l'action d'une force centrale de pôle l'origine $G$ et du type (6.52) avec

$$k = \frac{\mathcal{G}\mu m}{(1+\frac{m}{\mu})^2}. \tag{6.58}$$

Tout se passe comme si la formule (6.54) était applicable avec en $G$ la masse $\frac{\mu}{(1+\frac{m}{\mu})^2}$ et non $\mu$. Si $m$ est négligeable devant $\mu$, on a $G \approx O$ et $k \approx \mathcal{G}\mu m$. On retrouve le modèle de l'exemple 1.

► Mouvement par rapport au repère $\mathcal{R} = (O, (\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}))$.

On a

$$\mathbf{OM} = \mathbf{\Omega M} - \mathbf{\Omega O}.$$

Donc par dérivation

$$\left(\frac{d^2\mathbf{OM}}{dt^2}\right)_{\mathcal{R}_0} = \mathbf{\Gamma}(M/\mathcal{R}_0) - \mathbf{\Gamma}(O/\mathcal{R}_0). \tag{6.59}$$

Si $(x, y, z)$ sont les coordonnées dans $\mathcal{R}$ du point $M$, on a

$$\begin{aligned}
\left(\frac{d\mathbf{OM}}{dt}\right)_{\mathcal{R}_0} &= \frac{dx}{dt}\mathbf{e_1} + \frac{dy}{dt}\mathbf{e_2} + \frac{dz}{dt}\mathbf{e_3} \\
\left(\frac{d^2\mathbf{OM}}{dt^2}\right)_{\mathcal{R}_0} &= \frac{d^2x}{dt^2}\mathbf{e_1} + \frac{d^2y}{dt^2}\mathbf{e_2} + \frac{d^2z}{dt^2}\mathbf{e_3} \\
&= \mathbf{\Gamma}(M/\mathcal{R})
\end{aligned}$$

donc (6.59) s'écrit

$$\mathbf{\Gamma}(M/\mathcal{R}) = \mathbf{\Gamma}(M/\mathcal{R}_0) - \mathbf{\Gamma}(O/\mathcal{R}_0).$$

Comme $\mathcal{R}_0$ est galiléen, cette équation s'écrit encore

$$\begin{aligned}
\mathbf{\Gamma}(M/\mathcal{R}) &= \frac{1}{m}\mathbf{F}_M - \frac{1}{\mu}\mathbf{F}_O \\
&= \left(\frac{1}{m} + \frac{1}{\mu}\right)\mathbf{F}_M \\
&= \left(\frac{1}{m} + \frac{1}{\mu}\right)\left(-\frac{\mathcal{G}\mu m}{r^2}\mathbf{u}\right) \quad \text{d'après (6.55)} \\
&= -\frac{\mathcal{G}(m+\mu)}{r^2}\mathbf{u}.
\end{aligned}$$

On obtient donc :

$$m\,\mathbf{\Gamma}(M/\mathcal{R}) = -\frac{\mathcal{G}m(m+\mu)}{r^2}\mathbf{u}. \tag{6.60}$$

L'équation (6.60) du mouvement de $M$ dans le repère $\mathcal{R}$ est la même que si le repère $\mathcal{R}$ était galiléen et le point $M$ était soumis à une force centrale de pôle l'origine $O$ et du type (6.52) avec

$$k = \mathcal{G}(m+\mu)m. \tag{6.61}$$

Tout se passe comme si la formule (6.54) était applicable avec en $O$ la masse $m+\mu$ et non $\mu$.

**Énergie totale.**

On considère le champ de forces (6.52) avec l'hypothèse (6.53). $(x, y, z)$ désignent les coordonnées associées au repère $\mathcal{R}$.

Le champ de forces (6.52) dérive du *potentiel* $U = -\frac{k}{r}$, *i.e.*

$$\mathbf{F} = -\mathbf{grad}\, U$$

où $\mathbf{grad}\, U = \frac{\partial U}{\partial x}\mathbf{e_1} + \frac{\partial U}{\partial y}\mathbf{e_2} + \frac{\partial U}{\partial z}\mathbf{e_3}$. En effet, $r = \sqrt{x^2+y^2+z^2}$ donc $\frac{\partial U}{\partial x} = \frac{k}{r^2}\frac{x}{r}$, $\frac{\partial U}{\partial y} = \frac{k}{r^2}\frac{y}{r}$, $\frac{\partial U}{\partial z} = \frac{k}{r^2}\frac{z}{r}$ et

$$\mathbf{grad}\, U = \frac{k}{r^2}\left(\frac{x}{r}\mathbf{e_1} + \frac{y}{r}\mathbf{e_2} + \frac{z}{r}\mathbf{e_3}\right) = \frac{k}{r^2}\mathbf{u}$$

avec les notations de (6.52).

Considérons la *puissance*

$$\mathcal{P}_t = (\mathbf{F}|\mathbf{V}),$$

où $\mathbf{V} = \left(\frac{d\mathbf{OM}}{dt}\right)_{\mathcal{R}}$ désigne la vitesse de $M$. Comme le champ de forces dérive du potentiel $U$, on a :

$$\mathcal{P}_t = (-\mathbf{grad}\, U|\mathbf{V}) = -\left(\frac{\partial U}{\partial x}\frac{dx}{dt} + \frac{\partial U}{\partial y}\frac{dy}{dt} + \frac{\partial U}{\partial z}\frac{dz}{dt}\right).$$

Or

$$\frac{dU}{dt} = \frac{\partial U}{\partial t} + \frac{\partial U}{\partial x}\frac{dx}{dt} + \frac{\partial U}{\partial y}\frac{dy}{dt} + \frac{\partial U}{\partial z}\frac{dz}{dt}.$$

Comme le potentiel $U$ ne dépend que des coordonnées du point $M$ et non explicitement du temps, on a $\frac{\partial U}{\partial t} = 0$, donc

$$\mathcal{P}_t = (-\mathbf{grad}\, U|\mathbf{V}) = -\frac{dU}{dt}.$$

On dit alors que $U$ est *l'énergie potentielle de la masse ponctuelle $M$.*

*L'énergie cinétique de la masse ponctuelle $M$* est $T = \frac{1}{2}mv^2$ avec $v = \|\mathbf{V}\|$. On appelle $v$ la *vitesse numérique.*

L'équation du mouvement est, en notant $\mathbf{\Gamma} = \left(\frac{d^2\mathbf{OM}}{dt^2}\right)_{\mathcal{R}}$ l'accélération de $M$ par rapport à $\mathcal{R}$

$$\mathbf{F} = m\mathbf{\Gamma} \tag{6.62}$$

*i.e.*

$$\mathbf{\Gamma} = -\frac{k}{mr^2}\,\mathbf{u}. \tag{6.63}$$

D'après l'équation (6.62), on a :

$$\frac{dT}{dt} = \frac{d}{dt}\,\frac{1}{2}m(\mathbf{V}|\mathbf{V}) = m(\mathbf{V}|\mathbf{\Gamma}) = (\mathbf{V}|\mathbf{F}) = \mathcal{P}_t = -\frac{dU}{dt}$$

donc $T = -U + C^{ste}$ ou encore

$$E = T + U = C^{ste}. \tag{6.64}$$

$E = T + U$ est une constante du mouvement appelée *énergie totale de la masse ponctuelle $M$.* Une constante du mouvement est aussi appelée une *intégrale première.*

### 6.4.2 Trajectoire.

**Loi des aires.**

D'après (6.63)

$$\left(\frac{d}{dt}(\mathbf{OM} \wedge \mathbf{V})\right)_{\mathcal{R}} = \mathbf{OM} \wedge \mathbf{\Gamma} = 0$$

donc $\mathbf{OM} \wedge \mathbf{V}$ est un vecteur constant $\mathbf{C}$ :

$$\mathbf{OM} \wedge \mathbf{V} = \mathbf{C}. \tag{6.65}$$

Si le vecteur $\mathbf{C}$ est nul, on montre que le mouvement a lieu sur des demi-droites passant par le point $O$.

On suppose dans toute la suite $\mathbf{C} \neq 0$. D'après (6.65), le mouvement a alors lieu dans le plan $\Pi$ perpendiculaire au vecteur $\mathbf{C}$ et passant par $O$. Ce plan $\Pi$ étant fixe, on peut supposer que le repère galiléen $\mathcal{R}$ est tel que le plan $\Pi$ soit le plan de base orthonormée $(\mathbf{e_1}, \mathbf{e_2})$. On suppose $\Pi$ orienté par cette base. Alors $\mathbf{e_2}$ est le vecteur directement perpendiculaire à $\mathbf{e_1}$, *i.e.* obtenu par rotation d'angle $\frac{\pi}{2}$.

On considère le mouvement du point $M(t)$ sur un intervalle maximal $J$ contenant l'instant initial $t_0$. L'intervalle $J$ est alors nécessairement ouvert.

Soit $\theta = \widehat{(\mathbf{e_1}, \mathbf{u})}$ la détermination de l'angle $\widehat{(\mathbf{e_1}, \mathbf{u})}$ ($\mathbf{u} = \frac{\mathbf{OM}}{\|\mathbf{OM}\|}$) définie par continuité à partir des conditions initiales $(M(t_0), \theta(t_0))$. On a donc

$$\begin{aligned} \mathbf{u}(t) &= \cos\theta(t)\, \mathbf{e_1} + \sin\theta(t)\, \mathbf{e_2} & (6.66)\\ \mathbf{OM} &= r\mathbf{u} & (6.67) \end{aligned}$$

où $r = \|\mathbf{OM}\| > 0$. Le couple $(r, \theta)$ est un système de coordonnées polaires du point $M$, avec $r > 0$. $t \mapsto \theta(t)$ est une fonction de classe $C^1$ sur $J$. En notant par un $\dot{}$ la dérivation par rapport au temps $t$, (6.66) donne

$$\left(\frac{d\mathbf{u}}{dt}\right)_{\mathcal{R}} = \dot\theta\,(-\sin\theta\,\mathbf{e_1} + \cos\theta\,\mathbf{e_2}) = \dot\theta\,\mathbf{w} \tag{6.68}$$

où $\mathbf{w}$ est le vecteur directement perpendiculaire à $\mathbf{u}$. De même,

$$\left(\frac{d\mathbf{w}}{dt}\right)_{\mathcal{R}} = -\dot\theta\,\mathbf{u}.$$

Par dérivation par rapport au temps de (6.67) on obtient donc :

$$\begin{aligned} \mathbf{V} &= \dot r\,\mathbf{u} + r\dot\theta\,\mathbf{w} & (6.69)\\ \mathbf{\Gamma} &= (\ddot r - r\dot\theta^2)\,\mathbf{u} + (r\ddot\theta + 2\dot r\dot\theta)\,\mathbf{w}\,. & (6.70) \end{aligned}$$

Alors $\mathbf{OM} \wedge \mathbf{V} = r^2\dot\theta\;\mathbf{u} \wedge \mathbf{w}$, donc d'après (6.65), $r^2\dot\theta$ est une constante $C$ :

$$r^2\dot\theta = C. \tag{6.71}$$

$C$ est appelée la *constante des aires*. On a $|C| = \|\mathbf{C}\|$.

Cela implique en particulier que $\dot\theta$ ne s'annule pas donc garde un signe constant sur $J$. L'application $t \mapsto \theta(t)$ est une application dérivable strictement monotone de $J$ sur l'intervalle image $I = \theta(J)$. La fonction réciproque que nous noterons $\theta \mapsto t(\theta)$ de $I$ vers $J$ est aussi dérivable et l'on a $\frac{dt}{d\theta} = \frac{1}{\frac{d\theta}{dt}}$. Nous noterons encore $r$ la fonction $\theta \mapsto r(\theta) = r(t(\theta))$. Sa dérivée par rapport à $\theta$ est $r'(\theta) = \frac{dr}{d\theta} = \dot r\,\frac{dt}{d\theta}$.

**Formules de Binet.**

Par dérivation par rapport à $t$ de (6.71), on a après simplification par $r > 0$ :

$$r\ddot{\theta} + 2\dot{r}\dot{\theta} = 0. \tag{6.72}$$

D'autre part, $\dot{r} = r'(\theta)\dot{\theta}$ donne en utilisant (6.71) $\dot{r} = r'(\theta)\frac{C}{r^2}$, *i.e.*

$$\dot{r} = -C\alpha' \tag{6.73}$$

en notant $\alpha = \frac{1}{r}$ et $\alpha'$ sa dérivée par rapport à $\theta$. D'où

$$\ddot{r} = -C\alpha''\dot{\theta} = -C\alpha''\frac{C}{r^2} = -C^2\alpha^2\alpha'',$$

et comme $r\dot{\theta}^2 = (r^2\dot{\theta})^2\frac{1}{r^3} = C^2\alpha^3$,

$$\ddot{r} - r\dot{\theta}^2 = -C^2\alpha^2(\alpha'' + \alpha). \tag{6.74}$$

Les formules (6.69) et (6.70) pour $\mathbf{V}$ et $\mathbf{\Gamma}$ s'écrivent alors

$$\mathbf{V} = C(-\alpha'\,\mathbf{u} + \alpha\,\mathbf{w}) \tag{6.75}$$
$$\mathbf{\Gamma} = -C^2\alpha^2(\alpha'' + \alpha)\,\mathbf{u}. \tag{6.76}$$

Les formules (6.75) et (6.76) s'appellent les *formules de Binet.*

**Équation de la trajectoire.**

L'équation fondamentale (6.62) s'écrit en utilisant la formule de Binet (6.76) pour $\mathbf{\Gamma}$ :

$$mC^2\alpha^2(\alpha'' + \alpha) = k\alpha^2,$$

*i.e.*

$$\alpha'' + \alpha = \frac{k}{mC^2}. \tag{6.77}$$

Les solutions réelles $\theta \mapsto y(\theta)$ de l'équation différentielle sans second membre

$$y'' + y = 0$$

sont les fonctions

$$y(\theta) = \lambda\cos\theta + \mu\sin\theta$$

où $\lambda, \mu$ sont deux constantes réelles quelconques. En écrivant le nombre complexe $z = \lambda + i\mu$ sous la forme $z = Ae^{i\varphi}$ avec $A = \sqrt{\lambda^2 + \mu^2}$, ces solutions s'écrivent

$$y = A\cos(\theta - \varphi)$$

avec $A \geqslant 0$ et $\varphi \in \mathbb{R}$ des constantes d'intégration arbitraires. Les solutions réelles de l'équation avec second membre

$$y'' + y = \frac{k}{mC^2}$$

sont les fonctions

$$y = A\cos(\theta - \varphi) + \frac{k}{mC^2}$$

avec $A \geqslant 0$ et $\varphi \in \mathbb{R}$ des constantes d'intégration arbitraires.

Il existe donc des constantes $A \geqslant 0$ et $\varphi \in \mathbb{R}$ telles que

$$\alpha(\theta) = A\cos(\theta - \varphi) + \frac{k}{mC^2} \quad \forall\theta. \tag{6.78}$$

• Si $A = 0$, comme $\alpha > 0$ puisque $r > 0$, on a aussi $k > 0$ et la trajectoire du point $M$ est donc contenue dans le cercle de centre $O$ et de rayon $\frac{mC^2}{k}$.

• Si $A > 0$, on a

$$r = \frac{mC^2}{k} \frac{1}{1 + A\frac{mC^2}{k}\cos(\theta - \varphi)}$$

qui s'écrit encore

$$r = \frac{mC^2}{|k|} \frac{\varepsilon}{1 + \varepsilon e\cos(\theta - \varphi)} \tag{6.79}$$

où $\varepsilon = \frac{k}{|k|} = \pm 1$ représente le signe de $k$ et

$$e = A\frac{mC^2}{|k|}. \tag{6.80}$$

La trajectoire est donc contenue dans la courbe $\mathcal{C}$ dont une équation en coordonnées polaires est (6.79).

Si $k > 0$, on a $\varepsilon = 1$ et l'équation (6.79) est simplement

$$r = \frac{mC^2}{k} \frac{1}{1 + e\cos(\theta - \varphi)}. \tag{6.81}$$

D'après (6.18), $\mathcal{C}$ est une conique de foyer $O$, d'excentricité $e$ et dont l'axe focal a pour vecteur directeur $\cos\varphi\, \mathbf{e_1} + \sin\varphi\, \mathbf{e_2}$. Le paramètre de la conique $\mathcal{C}$ est

$$p = \frac{mC^2}{k}. \tag{6.82}$$

Si $k < 0$, on a $\varepsilon = -1$. L'équation (6.79) est

$$r = \frac{mC^2}{|k|} \frac{-1}{1 - e\cos(\theta - \varphi)}. \tag{6.83}$$

Dans le système de coordonnées polaires $(r_1, \theta_1)$ défini par $r_1 = -r, \theta_1 = \theta + \pi$, l'équation de $\mathcal{C}$ sera

$$r_1 = \frac{mC^2}{|k|} \frac{1}{1 + e\cos(\theta_1 - \varphi)}. \tag{6.84}$$

A nouveau d'après (6.18), on en déduit que $\mathcal{C}$ est une conique de foyer $O$, d'excentricité $e$ et dont l'axe focal a pour vecteur directeur $\cos\varphi\, \mathbf{e_1} + \sin\varphi\, \mathbf{e_2}$. Le paramètre de la conique $\mathcal{C}$ est

$$p = \frac{mC^2}{|k|}. \tag{6.85}$$

**Expression de l'excentricité $e$ avec l'énergie totale $E$.**

L'énergie totale est

$$E = T + U = \frac{1}{2}mv^2 - \frac{k}{r}.$$

L'équation (6.78) s'écrit par définition de $\alpha = \frac{1}{r}$

$$A\cos(\theta - \varphi) = \frac{1}{r} - \frac{k}{mC^2} \tag{6.86}$$

d'où par dérivation par rapport à $\theta$,

$$-A\sin(\theta - \varphi) = -\frac{r'}{r^2} = -\frac{\dot{r}}{r^2\dot{\theta}} = -\frac{\dot{r}}{C}. \tag{6.87}$$

On déduit de (6.86) et (6.87) :

$$A^2 = \frac{\dot{r}^2}{C^2} + \left(\frac{1}{r} - \frac{k}{mC^2}\right)^2, \tag{6.88}$$

donc d'après l'expression (6.80) de $e$

$$\begin{aligned} e^2 &= \frac{A^2m^2C^4}{k^2} \\ &= \dot{r}^2\frac{m^2C^2}{k^2} + \frac{m^2C^4}{k^2}\left(\frac{1}{r^2} - 2\frac{k}{mC^2r} + \frac{k^2}{m^2C^4}\right) \\ &= \dot{r}^2\frac{m^2C^2}{k^2} + \frac{m^2C^4}{k^2r^2} - 2\frac{mC^2}{kr} + 1 \\ &= 1 + \frac{mC^2}{k^2}\left(m\dot{r}^2 + \frac{mC^2}{r^2} - 2\frac{k}{r}\right) \\ &= 1 + \frac{mC^2}{k^2}\left(mv^2 - 2\frac{k}{r}\right) \end{aligned}$$

(car la vitesse numérique est $v = \sqrt{\dot{r}^2 + r^2\dot{\theta}^2} = \sqrt{\dot{r}^2 + \frac{C^2}{r^2}}$ ) ce qui donne finalement la formule

$$e^2 - 1 = 2\frac{mC^2}{k^2}E. \tag{6.89}$$

**Classification des trajectoires.**

**Théorème 6. 10.** *Soit $M(t)$ un point matériel de masse $m$ sur lequel s'exerce un champ de forces central en $\frac{1}{r^2}$ de pôle l'origine $O$, la force qui s'exerce sur le point $M$ étant donnée par l'équation (6.52) :*

$$\mathbf{F}_M = -\frac{k}{r^2}\,\mathbf{u}$$

*où $\mathbf{u} = \frac{\mathbf{OM}}{\|\mathbf{OM}\|}, r = \|\mathbf{OM}\|$, et $k$ est une constante $\neq 0$. On suppose les conditions initiales $M(t_0) = M_0, \mathbf{V}(t_0) = \mathbf{V}_0$, où $\mathbf{V}(t) = \left(\frac{dOM}{dt}\right)_{\mathcal{R}}$, le repère $\mathcal{R}$ galiléen (hypothèse (6.53)) et le mouvement étudié sur l'intervalle maximal $J = ]a, b[$ $(-\infty \leqslant a < t_0 < b \leqslant +\infty)$.*

*Soit $E$ l'énergie totale.*
*(i) On a*

$$E = \frac{1}{2} m v_0^2 - \frac{k}{r_0},$$

*avec $v_0 = \|\mathbf{V}_0\|$ et $r_0 = \|\mathbf{OM}_0\|$.*
*(ii) La trajectoire est portée par une conique $\mathcal{C}$ dont l'excentricité $e$ est donnée par la formule (6.89). On a :*

$$\begin{aligned} \mathcal{C} \text{ ellipse ou cercle } (0 \leqslant e < 1) &\Leftrightarrow E < 0 \\ \mathcal{C} \text{ parabole } (e = 1) &\Leftrightarrow E = 0 \\ \mathcal{C} \text{ hyperbole } (e > 1) &\Leftrightarrow E > 0 \end{aligned}$$

*(iii) On a $J = \mathbb{R}$. Si $\mathcal{C}$ est une ellipse ou un cercle, la trajectoire pour $t$ variant de $t_0$ à $+\infty$ est $\mathcal{C}$ entière, et le mouvement est périodique. Si $\mathcal{C}$ est une parabole, la trajectoire pour $t$ variant de $-\infty$ à $+\infty$ est $\mathcal{C}$ entière. Si $\mathcal{C}$ est une hyperbole, la trajectoire pour $t$ variant de $-\infty$ à $+\infty$ est une branche de l'hyperbole $\mathcal{C}$.*
*(iv) La trajectoire est un cercle si et seulement si*

$$\left\{ \begin{array}{rcl} \dot{r}_0 & = & 0 \\ v_0^2 & = & \frac{k}{m r_0} \end{array} \right. . \tag{6.90}$$

*Dans ce cas, le mouvement circulaire est uniforme : la vitesse angulaire $\dot{\theta}$ est constante, ainsi que la vitesse numérique $v$.*

**Démonstration.**

(i) On a vu que $E = \frac{1}{2} m v^2 - \frac{k}{r}$ avec $v = \|\mathbf{V}\|$ et $r = \|\mathbf{OM}\|$, et que $E$ est une constante du mouvement. On a donc $E(t) = E_0$ pour tout $t$, en notant $E_0 = E(t_0) = \frac{1}{2} m v_0^2 - \frac{k}{r_0}$.
(ii) Résulte directement des formules (6.79) et (6.89). On notera que par le choix fait pour les coordonnées polaires (6.67), on a $r > 0$ sur la trajectoire de $M$. Dans le cas $0 \leqslant e \leqslant 1$ on a dans (6.79) $1 + \varepsilon e \cos(\theta - \varphi) \geqslant 0$, donc l'équation (6.79) en un point de la trajectoire donne $k > 0$. Il y a *attraction*. Dans le cas $e > 1$ la condition $r > 0$ sur la trajectoire implique que $1 + \varepsilon e \cos(\theta - \varphi)$ est du signe de $k$ sur la trajectoire. Si $k > 0$ (*attraction*), on a l'équation (6.81) avec $r > 0$ sur la trajectoire. D'après la discussion dans l'Exemple (iii) qui suit le Théorème 6.9, la trajectoire est un arc de la branche de l'hyperbole relative au foyer $O$. Si $k < 0$ (*répulsion*), dans le système de coordonnées polaires $(r_1, \theta_1)$ défini par $r_1 = -r, \theta_1 = \theta + \pi$, l'équation de $\mathcal{C}$ est (6.84) avec $r_1 < 0$ sur la trajectoire. A nouveau d'après la discussion dans l'Exemple (iii), la trajectoire est un arc de la branche de l'hyperbole relative au foyer opposé .
(iii) D'après la loi des aires (6.71), l'application $t \mapsto \theta(t)$ est strictement monotone, croissante ou décroissante suivant le signe de la constante des aires $C$. Supposons $C > 0$, le cas $C < 0$ étant analogue. La fonction strictement croissante $t \mapsto \theta(t)$ possède une limite $\theta^* \leqslant +\infty$ quand $t \to b$. La loi des aires (6.71) $r^2 \dot{\theta} = C$ donne d'après (6.79) (avec $e = 0$ dans le cas d'un cercle)

$$\frac{dt}{d\theta} = \frac{r^2}{C} = \frac{m^2 C^3}{k^2} \frac{1}{(1 + \varepsilon e \cos{(\theta - \varphi)})^2}.$$

On a donc pour tout $t_1 \geqslant t_0$ en posant $\theta_0 = \theta(t_0)$ , $\theta_1 = \theta(t_1)$ :

$$t_1 - t_0 = \frac{m^2C^3}{k^2} \int_{\theta_0}^{\theta_1} \frac{d\theta}{(1+\varepsilon e \cos(\theta - \varphi))^2}. \tag{6.91}$$

Supposons $b < +\infty$. Si $\theta^* < +\infty$, $\alpha = \frac{1}{r}$ et $\alpha'$ possèdent une limite finie d'après (6.78), donc aussi $\mathbf{V}$ d'après la formule de Binet (6.75). Si la limite de $\alpha$ n'est pas nulle, $r$ a une limite finie $\neq 0$, donc $M$ possède une limite $\neq O$. Comme $M$ et $\mathbf{V}$ ont tous deux une limite quand $t \to b$, d'après un théorème sur les équations différentielles, le mouvement est prolongeable au delà de $b$, contrairement à la maximalité de $J$. Si la limite de $\alpha$ est nulle, $r$ tend vers $+\infty$ et la valeur $\theta^*$ annule $1 + \varepsilon e \cos(\theta - \varphi)$. Cela impose $e \geqslant 1$. Le développement de Taylor de la fonction $\theta \mapsto 1 + \varepsilon e \cos(\theta - \varphi)$ montre alors que $1 + \varepsilon e \cos(\theta - \varphi) \sim B(\theta - \theta^*)^j$ quand $\theta \to \theta^*$, où $B \in \mathbb{R}$ et $j \geqslant 1$. Donc l'intégrale $\int_{\theta_0}^{\theta^*} \frac{d\theta}{(1+\varepsilon e \cos(\theta-\varphi))^2}$ diverge, ce qui est en contradiction avec ce qu'on obtient par passage à la limite quand $t_1 \to b$ dans (6.91). Ce cas est donc impossible. Enfin, si $\theta^* = +\infty$, par passage à la limite dans (6.91) quand $t_1 \to b$, il vient $b - t_0 = \frac{m^2C^3}{k^2} \int_{\theta_0}^{+\infty} \frac{d\theta}{(1+\varepsilon e \cos(\theta-\varphi))^2}$. C'est impossible puisqu'il s'agit de l'intégrale d'une fonction périodique $> 0$ (avec en plus des pôles éventuels). Donc $b = +\infty$. On verrait de même que $a = -\infty$. Donc $J = \mathbb{R}$. Détaillons maintenant les 3 cas.

► Cas $0 \leqslant e < 1$. On a vu en (ii) que dans ce cas $k > 0$, *i.e.* $\varepsilon = 1$. Si $\theta^* < +\infty$, on aurait par passage à la limite dans (6.91) quand $t_1 \to +\infty$ :

$$+\infty = \frac{m^2C^3}{k^2} \int_{\theta_0}^{\theta^*} \frac{d\theta}{(1+e \cos(\theta - \varphi))^2}. \tag{6.92}$$

Mais ceci est impossible puisque le second membre est l'intégrale sur un compact d'une fonction continue. On en conclut que $\theta^* = +\infty$. Cela implique que la trajectoire est entièrement décrite une infinité de fois quand $t$ varie de $t_0$ à $+\infty$. Elle est en particulier périodique.

► Cas $e = 1$. On a vu en (ii) que dans ce cas aussi $k > 0$, *i.e.* $\varepsilon = 1$. La fonction $\theta \mapsto 1 + \cos(\theta - \varphi)$ s'annule pour $\theta - \varphi \equiv \pi \pmod{2\pi}$. Soit $N \in \mathbb{Z}$ tel que $\varphi + (2N-1)\pi < \theta_0 < \varphi + (2N+1)\pi$. Comme $r < +\infty$ pour tout $t$, $\theta([t_0, +\infty[)$ est un intervalle qui ne contient aucun réel de la forme $\varphi + (2k+1)\pi$ avec $k \in \mathbb{Z}$. Donc $\theta([t_0, +\infty[) \subset [\theta_0, \varphi + (2N+1)\pi[$. Si $\theta^* < \varphi + (2N+1)\pi$, on aurait comme précédemment l'équation (6.92) qui est impossible. Donc $\theta^* = \varphi + (2N+1)\pi$. Quand $t$ varie de $t_0$ à $+\infty$, l'arc de parabole correspondant à $\theta_0 \leqslant \theta < \varphi + (2N+1)\pi$ est entièrement décrit. Quand $t$ varie de $-\infty$ à $+\infty$, toute la parabole est décrite.

► Cas $e > 1$. La fonction $\theta \mapsto 1 + \varepsilon e \cos(\theta - \varphi)$ s'annule pour $\theta - \varphi \equiv \pm\psi \pmod{2\pi}$ où $\psi = \mathrm{Arccos}\,(-\frac{\varepsilon}{e})$. Cela correspond aux asymptotes de l'hyperbole. Il existe $N \in \mathbb{Z}$ tel que $-\psi + N\pi < \theta_0 - \varphi < \psi + N\pi$. Comme précédemment, $\theta([t_0, +\infty[) \subset [\theta_0,\ \varphi + \psi + N\pi[$ et $\theta^* = \varphi + \psi + N\pi$. On en conclut que quand $t$ varie de $t_0$ à $+\infty$, l'arc de la branche d'hyperbole correspondant à $\theta_0 \leqslant \theta < \varphi + \psi + N\pi$ est entièrement décrite. Quand $t$ varie de $-\infty$ à $+\infty$, toute la branche de l'hyperbole est décrite.

(iv) La trajectoire est un cercle si et seulement si la constante $A$ de (6.78) est nulle. Comme $A^2$ est donné par (6.88),

$$A = 0 \quad \Leftrightarrow \quad \left\{ \begin{array}{rcl} \dot{r}_0 & = & 0 \\ \frac{1}{r_0} & = & \frac{k}{mC^2} \end{array} \right.$$

$$\Leftrightarrow \quad \left\{ \begin{array}{rcl} \dot{r}_0 & = & 0 \\ \frac{C^2}{r_0^2} & = & \frac{k}{mr_0}\,. \end{array} \right.$$

Donc le mouvement est circulaire si et seulement si

$$\left\{\begin{array}{lcl} \dot{r}_0 & = & 0 \\ v_0^2 & = & \frac{k}{mr_0} \end{array}\right. \tag{6.93}$$

puisque d'après (6.69), la condition $\dot{r}_0 = 0$ implique $v_0^2 = r_0^2\dot{\theta}_0^2 = \frac{C^2}{r_0^2}$. Lorsque la condition (6.93) est vérifiée, puisque $A = 0$, on a de même *pour tout* $t$ d'après (6.88) $\left\{\begin{array}{lcl} \dot{r} & = & 0 \\ v^2 & = & \frac{k}{mr} \end{array}\right.$ et $v^2 = r^2\dot{\theta}^2$. En particulier, le mouvement circulaire est uniforme. □

**Remarque.** Chacun des deux cas $E < 0$ et $E = 0$ impose $k > 0$, donc *attraction*. Dans le cas $E > 0$, il peut y avoir attraction ou répulsion.

**Vitesse aréolaire.**

Soit $\mathcal{C}$ une courbe d'équation en coordonnées polaires $r = f(\theta)$ avec $f(\theta) \geqslant 0$. L'aire du secteur $S = \{\rho\,(\cos\varphi\ \mathbf{e_1} + \sin\varphi\ \mathbf{e_2})\ ;\ \rho \in [0, f(\varphi)];\ \theta_0 \leqslant \varphi \leqslant \theta_1\}$ est

$$\int\int_S dx\,dy = \int_{\theta_0}^{\theta_1} d\varphi \int_0^{f(\varphi)} \rho\,d\rho = \frac{1}{2}\int_{\theta_0}^{\theta_1} f(\varphi)^2 d\varphi. \tag{6.94}$$

Soit $t \mapsto \theta(t)$ une fonction dérivable croissante ou décroissante et considérons le point de la courbe $\mathcal{C}$ d'angle polaire $\theta(t)$. On appelle *aire balayée par le rayon vecteur entre les temps* $t_0$ *et* $t$ ($t_0 < t$) le nombre

$$\mathfrak{a}(t) = \frac{1}{2}\int_{\theta(t_0)}^{\theta(t)} f(\varphi)^2 d\varphi. \tag{6.95}$$

Si $\theta$ est croissante et $\theta(t) - \theta(t_0) \leqslant 2\pi$, c'est l'aire du secteur correspondant à $\theta_0 = \theta(t_0) \leqslant \varphi \leqslant \theta_1 = \theta(t)$. Si $\theta$ est décroissante, c'est une aire *négative* dont la valeur absolue est l'aire du secteur correspondant à $\theta_1 = \theta(t) \leqslant \varphi \leqslant \theta_0 = \theta(t_0)$. $\mathfrak{a}(t)$ est une fonction de $t$ dont la dérivée est

$$\frac{d}{dt}\,\mathfrak{a} = \frac{1}{2}f(\theta)^2\,\dot{\theta} = \frac{1}{2}r^2\dot{\theta}. \tag{6.96}$$

On appelle $\frac{1}{2}r^2\dot{\theta}$ la *vitesse aréolaire.* La vitesse aréolaire ne dépend bien entendu pas du point $t_0$ utilisé. Dire que le mouvement satisfait à la loi des aires (6.71) signifie que la vitesse aréolaire est constante.

### 6.4.3 Lois de Kepler.

On considère un cercle comme une ellipse dégénérée.

**Théorème 6. 11 (Lois de Kepler).** *1ère loi. Les trajectoires des planètes autour du Soleil sont des ellipses dont le Soleil est l'un des foyers.*
*2ème loi. La vitesse aréolaire de chaque planète est constante.*
*3ème loi. Le rapport entre le cube* $a^3$ *du grand axe de l'ellipse et le carré* $T^2$ *de la période du mouvement est le même pour toutes les planètes.*

**Démonstration.**

La planète est identifiée à un point $M$ de masse $m$. La force qui s'exerce sur la planète $M$ est la force due au champ de gravitation créé par le Soleil (on néglige les autres interactions) :

$$\mathbf{F}_M = -\frac{k}{r^2}\,\mathbf{u}$$

avec $k = \mathcal{G}\mu m$, $\mathcal{G}$ étant la constante de gravitation universelle et $\mu$ la masse du Soleil. D'après le Th. 6.10, la trajectoire de $M$ est une ellipse, une hyperbole ou une parabole. Dans les deux derniers cas, la planète sortirait du système solaire. D'où la 1ère loi.

Le mouvement satisfait à la loi des aires, et donc d'après (6.96) la vitesse aréolaire est constante, égale à $\frac{C}{2}$, où $C$ est la constante des aires. D'où la 2ème loi.

Enfin, soit $T$ la période du mouvement. Pour $t_0$ fixé, l'ellipse est décrite par la planète quand $t$ varie de $t_0$ à $t_0 + T$. On a vu que, si $r = f(\theta)$ est l'équation en coordonnées polaires de l'ellipse, l'aire balayée par le rayon vecteur entre les temps $t_0$ et $t$ est une fonction dérivable de $t$, et sa dérivée $\frac{d}{dt}\mathfrak{a}$ est la vitesse aréolaire $\frac{1}{2}r^2\dot{\theta} = \frac{C}{2}$. Elle est du signe (constant) de $\dot{\theta}$. L'aire de l'ellipse entière est

$$|\mathfrak{a}(t_0+T) - \mathfrak{a}(t_0)| = \left|\int_{t_0}^{t_0+T} \frac{d}{dt}\mathfrak{a}\,dt\right| = \left|\frac{1}{2}\int_{t_0}^{t_0+T} C\,dt\right| = \frac{1}{2}\,|C|\,T.$$

Or l'aire de l'ellipse est $\pi ab$. En effet, l'ellipse se déduit du cercle principal par l'affinité orthogonale (6.2), dont le jacobien est $\begin{vmatrix} 1 & 0 \\ 0 & \frac{b}{a} \end{vmatrix} = \frac{b}{a}$. Par cette affinité, les aires sont ainsi multipliées par $\frac{b}{a}$. Le cercle principal a pour aire $\pi a^2$, donc l'ellipse a pour aire $\pi ab$. Les deux expressions calculées pour l'aire de l'ellipse donnent alors

$$\frac{1}{2}\,|C|\,T = \pi ab. \tag{6.97}$$

Par ailleurs, le paramètre $p$ de la trajectoire est d'après (6.85)

$$p = \frac{mC^2}{k} = \frac{C^2}{\mathcal{G}\mu}.$$

Mais on sait que le paramètre d'une ellipse est $p = \frac{b^2}{a}$. Donc

$$\frac{b^2}{a} = \frac{C^2}{\mathcal{G}\mu}. \tag{6.98}$$

L'élimination de $C$ entre les équations (6.97) et (6.98) donne

$$\frac{b^2\mathcal{G}\mu}{a} = \frac{4\pi^2a^2b^2}{T^2}, \tag{6.99}$$

*i.e.*

$$\frac{a^3}{T^2} = \frac{\mathcal{G}\mu}{4\pi^2}. \tag{6.100}$$

D'où la 3ème loi. □

**Remarques.** (i) Les lois de Kepler (Johannes Kepler : 1571-1630) sont antérieures à la formulation de la loi de Newton (Isaac Newton : 1642-1727), et ont été formulées à partir des observations de l'astronome Tycho Brahe (1545-1601).

(ii) La 3ème loi de Kepler est une approximation : on a en effet négligé l'action de la planète sur le Soleil et fait l'hypothèse (6.53). On sait que pour tenir compte de cette action il suffit de prendre pour $k$ la valeur définie par (6.61), *i.e.* remplacer $\mu$ par $\mu + m$. Cela n'influe pas sur les 2 premières lois. Mais cela conduit dans la 3ème loi à la valeur

$$\frac{a^3}{T^2} = \frac{\mathcal{G}(\mu + m)}{4\pi^2} \tag{6.101}$$

qui dépend de la planète en question.

**Quelques données planétaires.**

Le tableau suivant donne pour chacune des planètes : $\Delta$ = diamètre en km ; $m$ = masse en $10^{24}$ kg ; $a$ = grand axe en $10^6$ km; $T$ = période en années; rapport $\frac{a^3}{T^2}$ en $10^{24}(\text{km})^3/(\text{année})^2$; rapport $\frac{a^3}{T^2}$ en $10^{18}$ unités SI; $e$ = excentricité . Le coefficient multiplicateur (basé sur 1 année = $31.56 \cdot 10^6$ secondes) faisant passer du rapport $\frac{a^3}{T^2}$ en $10^{24}(\text{km})^3/(\text{année})^2$ au rapport en $10^{18}$ unités SI est $\frac{10^{24}}{10^{18}} \frac{10^9}{(31.56 \cdot 10^6)^2} = \frac{10^3}{(31.56)^2} = 1.00398 \approx 1.004$ .

| Planète | $\Delta$ | $m$ | $a$ | $T$ | $\frac{a^3}{T^2}$ | $\frac{a^3}{T^2}$ $10^{18}$ SI | $e$ |
|---|---|---|---|---|---|---|---|
| Mercure | 4879 | 0.33 | 57.9 | 0.241 | 3.341 | 3.354 | 0.206 |
| Venus | 12104 | 4.87 | 108.2 | 0.615 | 3.349 | 3.362 | 0.007 |
| Terre | 12756 | 5.97 | 149.6 | 1 | 3.341 | 3.354 | 0.017 |
| Mars | 6794 | 0.64 | 227.9 | 1.88 | 3.349 | 3.362 | 0.093 |
| Jupiter | 142984 | 1898.6 | 778.3 | 11.86 | 3.351 | 3.364 | 0.049 |
| Saturne | 120536 | 568.46 | 1427 | 29.5 | 3.339 | 3.352 | 0.056 |
| Uranus | 51118 | 86.83 | 2870 | 84 | 3.350 | 3.363 | 0.046 |
| Neptune | 49528 | 102.43 | 4495 | 165 | 3.336 | 3.349 | 0.011 |
| Pluton | 2390 | 0.012 | 5900 | 248 | 3.339 | 3.352 | 0.244 |

La masse $\mu$ du Soleil étant approximativement $\mu \approx 1.99 \cdot 10^{30}$ kg, on a en unités SI $\frac{\mu\mathcal{G}}{4\pi^2} \approx \frac{1.99 \cdot 10^{30} \cdot 6.67 \cdot 10^{-11}}{4\pi^2} \approx 3.36 \cdot 10^{18}$. Les données du tableau sont en accord avec la 3ème loi de Kepler. Le tableau a été établi d'après les données de [13]. On peut aussi utiliser le calculateur astro-physique [14].

## 6.5 Exercices.

Dans les exercices 6.1 à 6.7, on considère le plan affine euclidien rapporté à un repère orthonormé $(O, (\mathbf{e_1}, \mathbf{e_2}))$ et muni de l'orientation associée.

**Exercice 6.1.**

Quelle est la nature de la courbe d'équation

$$x^2 + 2xy + y^2 - 4y + 3 = 0\ ?$$

La dessiner et préciser les coordonnées $(x_\Omega, y_\Omega)$ de son sommet $\Omega$, le foyer et la directrice.

**Indication.**

Voir Fig. 6.16. On a $x^2 + 2xy + y^2 = {}^t\binom{x}{y} A \binom{x}{y}$ avec $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. Les vecteurs $\mathbf{f_1} = \frac{1}{\sqrt{2}}\binom{1}{-1}$ et $\mathbf{f_2} = \frac{1}{\sqrt{2}}\binom{1}{1}$ forment une base orthonormée directe, et $\mathbf{f_1}$

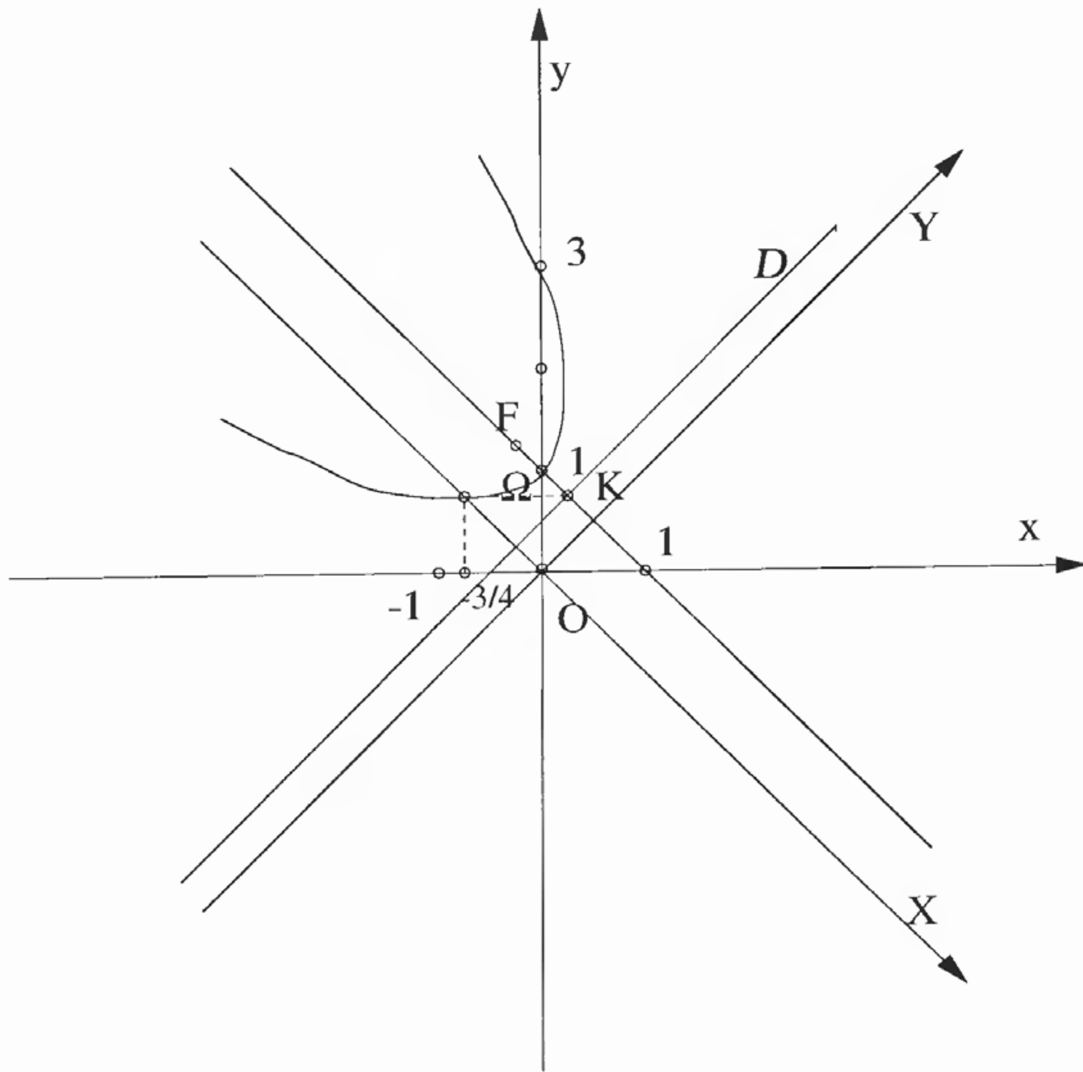



FIG. 6.16: *Exercice 6.1*

(*resp.* $\mathbf{f_2}$) est vecteur propre de $A$ pour la valeur propre 0 (*resp.* 2). Un vecteur $\mathbf{v} = \binom{x}{y}$ a pour composantes dans cette nouvelle base $X, Y$ tels que $\binom{x}{y} = P\binom{X}{Y}$ avec $P = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$. La courbe a pour équation dans cette base :

$$2Y^2 - \frac{4}{\sqrt{2}}(-X+Y) + 3 = 0,$$

qui s'écrit

$$(Y - \frac{\sqrt{2}}{2})^2 + X\sqrt{2} + 1 = 0,$$

*i.e.*

$$(Y - \frac{\sqrt{2}}{2})^2 = -\sqrt{2}(X + \frac{\sqrt{2}}{2}).$$

C'est une parabole de sommet $\Omega$ : $X_\Omega = -\frac{\sqrt{2}}{2}, Y_\Omega = \frac{\sqrt{2}}{2}$, de paramètre $p = \frac{\sqrt{2}}{2}$ tournée vers les $X < 0$. On a $x_\Omega = 0, y_\Omega = 1$. Le foyer $F$ est tel que $\mathbf{OF} = \mathbf{O\Omega} - \frac{p}{2}\mathbf{f_1}$ donc $x_F = -\frac{1}{4}, y_F = \frac{5}{4}$. La directrice $\mathcal{D}$ a pour vecteur directeur $\mathbf{f_2}$ et passe par le point $K$ : $x_K = \frac{1}{4}, y_K = \frac{3}{4}$.

**Exercice 6.2.**

Quelle est la nature de la courbe d'équation

$$x^2 - 3xy + 2y^2 + 2x - 4y + 1 = 0 \ ?$$

Préciser les coordonnées $(x_\Omega, y_\Omega)$ de son centre $\Omega$ , l'excentricité et les asymptotes.

**Indication.**

Voir Fig. 6.17. On a $x^2 - 3xy + 2y^2 = {}^t\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) A \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ avec $A = \begin{pmatrix} 1 & -\frac{3}{2} \\ -\frac{3}{2} & 2 \end{pmatrix}$. Les vecteurs $\mathbf{f_1} = \frac{1}{\sqrt{2}} \begin{pmatrix} \frac{3}{\sqrt{10-\sqrt{10}}} \\ \frac{\sqrt{10}-1}{\sqrt{10-\sqrt{10}}} \end{pmatrix}$ et $\mathbf{f_2} = \frac{1}{\sqrt{2}} \begin{pmatrix} \frac{-3}{\sqrt{10+\sqrt{10}}} \\ \frac{\sqrt{10}+1}{\sqrt{10+\sqrt{10}}} \end{pmatrix}$ forment une base orthonormée directe, et $\mathbf{f_1}$ (*resp.* $\mathbf{f_2}$) est vecteur propre de $A$ pour la valeur propre $-\frac{\sqrt{10}-3}{2}$ (*resp.* $\frac{\sqrt{10}+3}{2}$). Un vecteur $\mathbf{v} = \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ a pour composantes dans cette nouvelle base $X, Y$ tels que $\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = P \left(\begin{smallmatrix} X \\ Y \end{smallmatrix}\right)$ avec $P = \frac{1}{\sqrt{2}} \begin{pmatrix} \frac{3}{\sqrt{10-\sqrt{10}}} & \frac{-3}{\sqrt{10+\sqrt{10}}} \\ \frac{\sqrt{10}-1}{\sqrt{10-\sqrt{10}}} & \frac{\sqrt{10}+1}{\sqrt{10+\sqrt{10}}} \end{pmatrix}$. La courbe a pour équation dans cette base :

$$-\frac{\sqrt{10}-3}{2}X^2 + \frac{\sqrt{10}+3}{2}Y^2 + \frac{\sqrt{2}(5-2\sqrt{10})}{\sqrt{10-\sqrt{10}}}X - \frac{\sqrt{2}(5+2\sqrt{10})}{\sqrt{10+\sqrt{10}}}Y + 1 = 0. \tag{6.102}$$

Posons

$$X_\Omega = \frac{(5-2\sqrt{10})\sqrt{2}}{(\sqrt{10}-3)\sqrt{10-\sqrt{10}}}$$

$$Y_\Omega = \frac{(5+2\sqrt{10})\sqrt{2}}{(\sqrt{10}+3)\sqrt{10+\sqrt{10}}}.$$

Comme

$$\frac{(5-2\sqrt{10})^2}{(\sqrt{10}-3)(10-\sqrt{10})} - \frac{(5+2\sqrt{10})^2}{(\sqrt{10}+3)(10+\sqrt{10})} = 0,$$

l'équation (6.102) s'écrit :

$$-\frac{\sqrt{10}-3}{2}(X - X_\Omega)^2 + \frac{\sqrt{10}+3}{2}(Y - Y_\Omega)^2 + 1 = 0.$$

*i.e.*

$$\frac{\sqrt{10}-3}{2}(X - X_\Omega)^2 - \frac{\sqrt{10}+3}{2}(Y - Y_\Omega)^2 = 1. \tag{6.103}$$

L'équation (6.103) est l'équation dans la base $(\mathbf{f_1}, \mathbf{f_2})$ d'une hyperbole de centre $\Omega$ de coordonnées $(X_\Omega, Y_\Omega)$, de distance focale $c = 2\sqrt[4]{10}$ , excentricité $e = \sqrt{2(10-3\sqrt{10})} \approx 1.013$. Les asymptotes ont pour équation $Y - Y_\Omega = \pm(\sqrt{10}-3)(X - X_\Omega)$. Le calcul de $x_\Omega, y_\Omega$ et de l'équation en $x, y$ des asymptotes est un peu fastidieux. On trouve d'abord $x_\Omega = -4, y_\Omega = -2$. Ensuite l'asymptote $Y - Y_\Omega = (\sqrt{10}-3)(X - X_\Omega)$ a pour vecteur directeur le vecteur $\mathbf{v}$ de composantes $\left(\begin{smallmatrix} 1 \\ \sqrt{10}-3 \end{smallmatrix}\right)$ dans la base $(\mathbf{f_1}, \mathbf{f_2})$; les coordonnées de $\mathbf{v}$ dans la base

canonique sont alors

$$\begin{aligned} u &= \frac{3\sqrt{\sqrt{10}}}{\sqrt{2}\sqrt{90}}\left(\sqrt{\sqrt{10}+1}-(\sqrt{10}-3)\sqrt{\sqrt{10}-1}\right) \\ v &= \frac{\sqrt{\sqrt{10}}}{\sqrt{2}\sqrt{90}}\left(\sqrt{\sqrt{10}+1}(\sqrt{10}-1)+(\sqrt{10}-3)(\sqrt{10}+1)\sqrt{\sqrt{10}-1}\right) \\ &= \frac{3\sqrt{\sqrt{10}}}{\sqrt{2}\sqrt{90}}\left(\sqrt{\sqrt{10}-1}+(\sqrt{10}-3)\sqrt{\sqrt{10}+1}\right). \end{aligned}$$

Les deux nombres

$$\sqrt{\sqrt{10}+1}-(\sqrt{10}-3)\sqrt{\sqrt{10}-1}$$

et

$$\sqrt{\sqrt{10}-1}+(\sqrt{10}-3)\sqrt{\sqrt{10}+1}$$

sont $> 0$ et par élévation au carré on voit qu'ils sont égaux. Donc $u = v$. Comme l'asymptote passe par $\Omega$, on en déduit qu'elle a pour équation

$$y = x + 2.$$

On obtient de même pour l'autre asymptote l'équation $y = \frac{1}{2}x$.

On peut éviter ces calculs fastidieux. Pour le calcul de $x_\Omega$ et $y_\Omega$, on sait qu'ils sont les solutions du système

$$\begin{cases} \frac{\partial f}{\partial x} &= 0 \\ \frac{\partial f}{\partial y} &= 0 \end{cases} \tag{6.104}$$

en notant

$$f(x,y) = x^2 - 3xy + 2y^2 + 2x - 4y + 1.$$

Cela donne immédiatement $x_\Omega = -4, y_\Omega = -2$. Ensuite pour les asymptotes, sachant que la courbe est une hyperbole, les directions asymptotiques sont obtenues comme suit : pour $x \neq 0$, l'équation $f(x,y) = 0$ s'écrit en divisant par $x^2$

$$1 - 3\frac{y}{x} + 2\frac{y^2}{x^2} + \frac{2}{x} - \frac{4y}{x^2} + \frac{1}{x^2} = 0. \tag{6.105}$$

Or une asymptote non verticale a une pente $t$, et si $(x, y(x))$ est un point de l'hyperbole qui tend vers cette asymptote quand $x \to +\infty$, on aura $\frac{y}{x} \to t$ donc :

$$1 - 3t + 2t^2 = 0. \tag{6.106}$$

On en déduit que les directions asymptotiques sont $t = 1$ et $t = \frac{1}{2}$. Les équations des deux asymptotes s'en déduisent immédiatement. On notera que la courbe est tangente à l'axe des $x$ au point $x = -1, y = 0$ qui n'est pas sur l'axe focal.

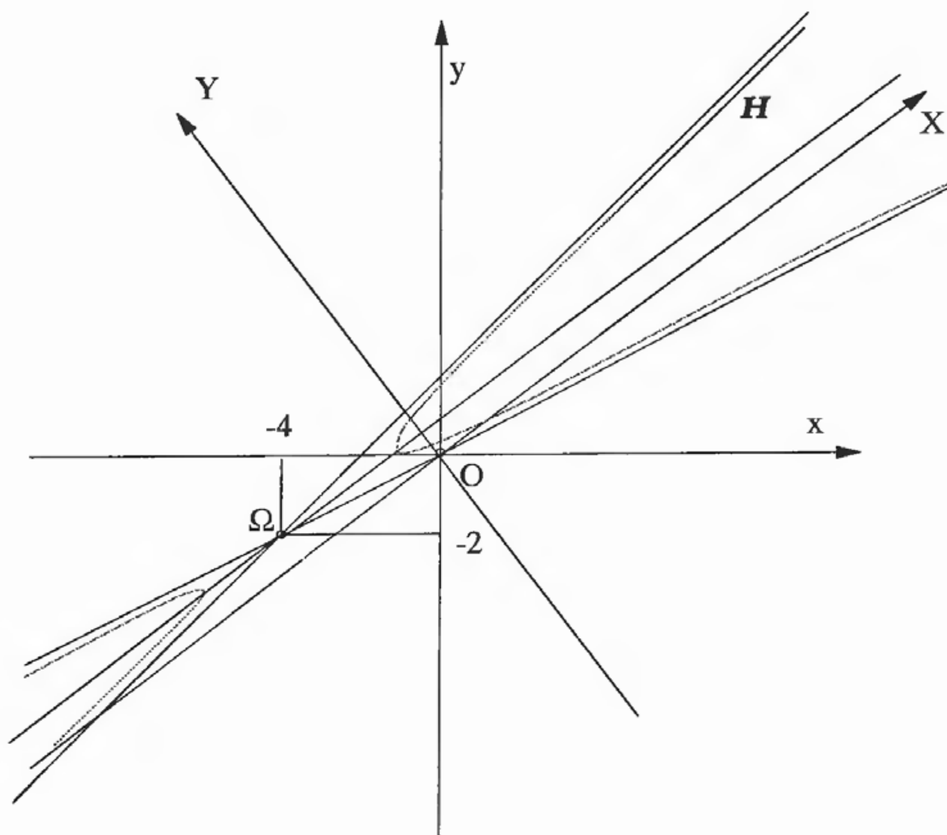



FIG. 6.17: *Exercice 6.2*

**Exercice 6.3.**

Quelle est la nature de la courbe d'équation

$$x^2 + 10xy + y^2 + 2x - y = 3\ ?$$

Préciser les coordonnées $(x_\Omega, y_\Omega)$ de son centre $\Omega$, l'excentricité, un foyer et la directrice associée, et éventuellement les asymptotes.

**Indication.**

Voir Fig. 6.18. On a $x^2 + 10xy + y^2 = {}^t\begin{pmatrix} x \\ y \end{pmatrix} A \begin{pmatrix} x \\ y \end{pmatrix}$ avec $A = \begin{pmatrix} 1 & 5 \\ 5 & 1 \end{pmatrix}$. Les vecteurs $\mathbf{f_1} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et $\mathbf{f_2} = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ forment une base orthonormée directe, et $\mathbf{f_1}$ (*resp.* $\mathbf{f_2}$) est vecteur propre de $A$ pour la valeur propre 6 (*resp.* $-4$). Un vecteur $\mathrm{v} = \begin{pmatrix} x \\ y \end{pmatrix}$ a pour composantes dans cette nouvelle base $X, Y$ tels que $\begin{pmatrix} x \\ y \end{pmatrix} = P\begin{pmatrix} X \\ Y \end{pmatrix}$ avec $P = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$. La courbe a pour équation dans cette base :

$$6X^2 - 4Y^2 + \frac{2}{\sqrt{2}}(X - Y) - \frac{1}{\sqrt{2}}(X + Y) = 3,$$

qui s'écrit

$$6(X + \frac{1}{12\sqrt{2}})^2 - 4(Y + \frac{3}{8\sqrt{2}})^2 - \frac{1}{48} + \frac{9}{32} = 3$$

ou encore

$$6(X - X_\Omega)^2 - 4(Y - Y_\Omega)^2 = \frac{263}{96} \tag{6.107}$$

en posant $X_\Omega = -\frac{1}{12\sqrt{2}}$ et $Y_\Omega = -\frac{3}{8\sqrt{2}}$. On a alors $x_\Omega = \frac{7}{48} \approx 0.15$, $y_\Omega = -\frac{11}{48} \approx -0.23$. L'équation (6.107) est l'équation dans la base $(\mathbf{f_1}, \mathbf{f_2})$ d'une hyperbole de centre $\Omega$ de coordonnées $(X_\Omega, Y_\Omega)$, grand axe $a = \frac{\sqrt{263}}{24} \approx 0.68$, petit axe

$b = \frac{\sqrt{263}}{8\sqrt{6}} \approx 0.83$. La distance focale $c$ est donnée par $c^2 = a^2 + b^2 = \frac{263}{96}(\frac{1}{6} + \frac{1}{4}) = \frac{263}{96}\frac{5}{12} \approx 1.14$ donc $c = \frac{\sqrt{263}\sqrt{10}}{48} \approx 1.07$. L'excentricité est $e = \frac{\sqrt{10}}{2} \approx 1.58$. L'un des foyers $F$ est défini par $\Omega\mathbf{F} = c\mathbf{f_1} = \frac{\sqrt{263}\sqrt{5}}{48}\binom{1}{1}$, d'où $x_F \approx 0.90$ et $y_F \approx 0.53$. Le point $K$ d'intersection avec l'axe focal de la directrice relative au foyer $F$ est donné par $\Omega\mathbf{K} = \frac{a}{e}\mathbf{f_1} = \frac{\sqrt{263}}{24\sqrt{5}}\binom{1}{1}$, d'où $x_K \approx 0.45$ et $y_K \approx 0.07$. La directrice a donc pour équation en $x, y$ dans le repère $(O, (\mathbf{e_1}, \mathbf{e_2}))$ :

$$y - y_K = -(x - x_K),$$

soit

$$y = -x + x_K + y_K = -x + \frac{\sqrt{263}}{12\sqrt{5}} - \frac{1}{12} \approx -x + 0.52\,.$$

Les asymptotes ont pour équation en $X, Y$ dans le repère $(O, (\mathbf{f_1}, \mathbf{f_2}))$ $Y - Y_\Omega = \pm\sqrt{\frac{3}{2}}(X - X_\Omega)$. Le calcul de l'équation en $x, y$ des asymptotes dans le repère $(O, (\mathbf{e_1}, \mathbf{e_2}))$ est fait comme dans l'exercice 6.2, équation (6.106). Les directions asymptotiques sont données par

$$1 + 10t + t^2 = 0 \tag{6.108}$$

et sont donc $t_1 = -5 + 2\sqrt{6} \approx -0.1$ et $t_2 = -5 - 2\sqrt{6} =\approx -9.9$. Les équations des deux asymptotes s'en déduisent immédiatement :

$$y - y_\Omega = (-5 \pm 2\sqrt{6})(x - x_\Omega).$$

Les points d'intersection de l'hyperbole avec l'axe des $y$ sont donnés par l'équation $y^2 - y = 3$, dont les racines sont $\frac{1+\sqrt{13}}{2} \approx 2.3$ et $\frac{1-\sqrt{13}}{2} \approx -1.3$. Les points d'intersection de l'hyperbole avec l'axe des $x$ sont donnés par l'équation $x^2 + 2x = 3$, dont les racines sont $-3$ et $1$.

**Exercice 6.4.**

Quelle est la nature de la courbe d'équation

$$4x^2 + 24xy + 11y^2 + 25x + 20 = 0\ ?$$

Préciser les coordonnées $(x_\Omega, y_\Omega)$ de son centre $\Omega$, l'excentricité, un foyer et la directrice associée, et éventuellement les asymptotes.

**Indication.**

Voir Fig. 6.19. On a $4x^2 + 24xy + 11y^2 = {}^t\binom{x}{y} A \binom{x}{y}$ avec $A = \begin{pmatrix} 4 & 12 \\ 12 & 11 \end{pmatrix}$. Les vecteurs $\mathbf{f_1} = \frac{1}{5}\binom{4}{-3}$ et $\mathbf{f_2} = \frac{1}{5}\binom{3}{4}$ forment une base orthonormée directe, et $\mathbf{f_1}$ (*resp.* $\mathbf{f_2}$) est vecteur propre de $A$ pour la valeur propre $-5$ (*resp.* 20). Un vecteur $\mathbf{v} = \binom{x}{y}$ a pour composantes dans cette nouvelle base $X, Y$ tels que $\binom{x}{y} = P\binom{X}{Y}$ avec $P = \frac{1}{5}\begin{pmatrix} 4 & 3 \\ -3 & 4 \end{pmatrix}$. La courbe a pour équation dans cette base :

$$-5X^2 + 20Y^2 + 5(4X + 3Y) + 20 = 0,$$

qui s'écrit

$$-(X - 2)^2 + 4(Y + \frac{3}{8})^2 + 4 - \frac{9}{16} + 4 = 0$$

ou encore

$$(X - X_\Omega)^2 - 4(Y - Y_\Omega)^2 = \frac{119}{16} \tag{6.109}$$

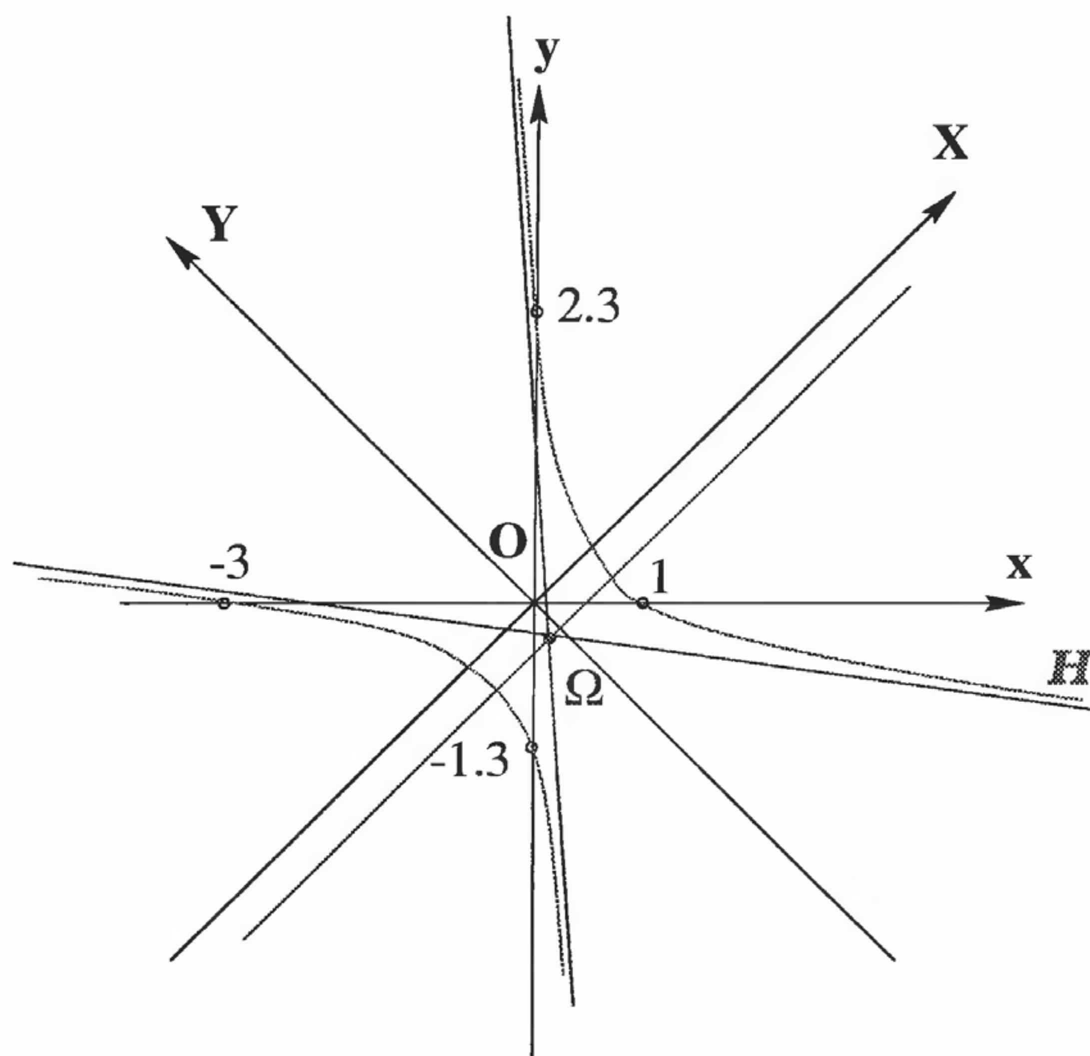



FIG. 6.18: *Exercice 6.3*

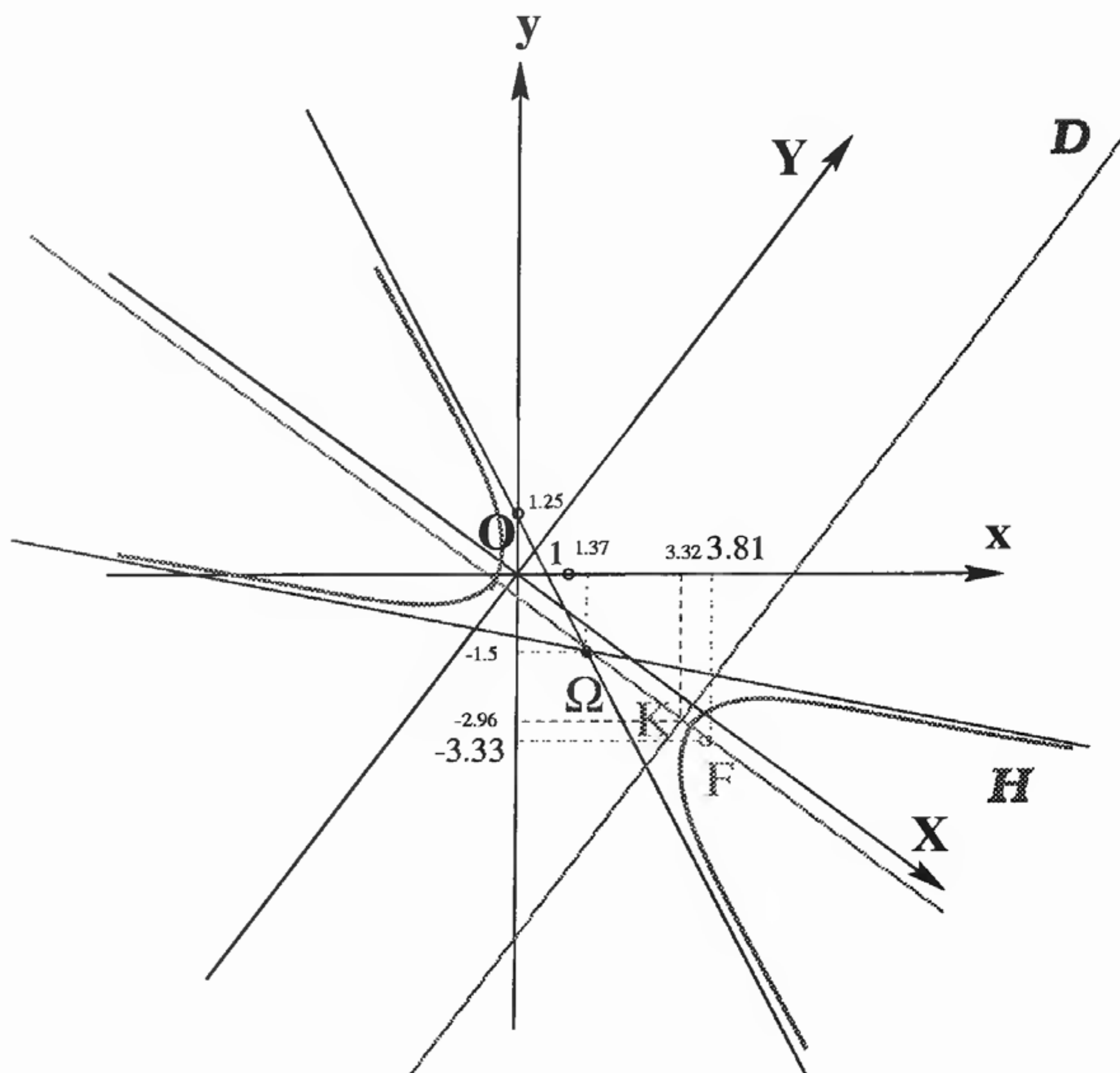



FIG. 6.19: *Exercice 6.4*

en posant $X_\Omega = 2$ et $Y_\Omega = -\frac{3}{8} \approx -0.37$. On a alors $x_\Omega = \frac{11}{8} \approx 1.37$, $y_\Omega = -\frac{3}{2}$. L'équation (6.109) est l'équation dans la base $(\mathbf{f_1}, \mathbf{f_2})$ d'une hyperbole de centre $\Omega$ de coordonnées $(X_\Omega, Y_\Omega)$, grand axe $a = \frac{\sqrt{119}}{4} \approx 2.73$, petit axe $b = \frac{a}{2} = \frac{\sqrt{119}}{8} \approx 1.36$. La distance focale c est donnée par $c^2 = a^2 + b^2 = \frac{595}{64}$, soit $c = \frac{\sqrt{595}}{8} \approx 3.05$. L'excentricité est $e = \frac{\sqrt{5}}{2} \approx 1.12$. L'un des foyers $F$ est défini par $\mathbf{\Omega F} = c\,\mathbf{f_1} = \frac{\sqrt{595}}{40}\begin{pmatrix} 4 \\ -3 \end{pmatrix}$, d'où $x_F = \frac{4\sqrt{595}+55}{40} \approx 3.81$ et $y_F = -\frac{3(\sqrt{595}+20)}{20} \approx -3.33$. Le point $K$ d'intersection avec l'axe focal de la directrice relative au foyer $F$ est donné par $\mathbf{\Omega K} = \frac{a}{e}\,\mathbf{f_1} = \frac{\sqrt{119}}{2\sqrt{5}}\,\frac{1}{5}\begin{pmatrix} 4 \\ -3 \end{pmatrix}$, d'où $x_K \approx 3.32$ et $y_K \approx -2.96$. La directrice a donc pour équation en $x, y$ dans le repère $(O, (\mathbf{e_1}, \mathbf{e_2}))$ :

$$y - y_K = \frac{4}{3}(x - x_K),$$

soit

$$y = \frac{4}{3}x - \frac{20 + \sqrt{595}}{6}.$$

Les asymptotes ont pour équation en $X, Y$ dans le repère $(O, (\mathbf{f_1}, \mathbf{f_2}))$

$$Y - Y_\Omega = \pm\frac{1}{2}(X - X_\Omega).$$

Le calcul de l'équation en $x, y$ des asymptotes dans le repère $(O, (\mathbf{e_1}, \mathbf{e_2}))$ est fait comme dans l'exercice 6.2, équation (6.106). Les directions asymptotiques sont données par

$$4 + 24t + 11t^2 = 0 \tag{6.110}$$

et sont donc $t_1 = -\frac{2}{11} \approx -0.18$ et $t_2 = -2$. Les équations des deux asymptotes s'en déduisent immédiatement :

$$\begin{aligned} y - y_\Omega &= -\frac{2}{11}(x - x_\Omega), \\ y - y_\Omega &= -2(x - x_\Omega), \end{aligned}$$

soit

$$\begin{aligned} y &= -\frac{2}{11}x - \frac{5}{4}, \\ y &= -2x + \frac{5}{4}. \end{aligned}$$

Il n'y a pas de point d'intersection de l'hyperbole avec l'axe des $y$. Les points d'intersection de l'hyperbole avec l'axe des $x$ sont donnés par l'équation $4x^2 + 25x + 20 = 0$, dont les racines sont $\frac{-25 \pm \sqrt{305}}{8}$, soit approximativement $-0.94$ et $-5.3$.

**Exercice 6.5.**

Quelle est la nature de la courbe d'équation

$$31x^2 - 24xy + 21y^2 + 4x + 6y = 25\ ?$$

Préciser les coordonnées $(x_\Omega, y_\Omega)$ de son centre $\Omega$, l'excentricité, un foyer et la directrice associée.

**Indication.**

Voir Fig. 6.20. On a $31x^2 - 24xy + 21y^2 = {}^t\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) A \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ avec $A = \left(\begin{smallmatrix} 31 & -12 \\ -12 & 21 \end{smallmatrix}\right)$. Les vecteurs $\mathbf{f_1} = \frac{1}{\sqrt{13}}\left(\begin{smallmatrix} 2 \\ 3 \end{smallmatrix}\right)$ et $\mathbf{f_2} = \frac{1}{\sqrt{13}}\left(\begin{smallmatrix} -3 \\ 2 \end{smallmatrix}\right)$ forment une base orthonormée directe, et $\mathbf{f_1}$ (*resp.* $\mathbf{f_2}$) est vecteur propre de $A$ pour la valeur propre 13 (*resp.* 39). Un vecteur $\mathbf{v} = \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ a pour composantes dans cette nouvelle base $X, Y$ tels que $\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = P \left(\begin{smallmatrix} X \\ Y \end{smallmatrix}\right)$ avec $P = \frac{1}{\sqrt{13}}\left(\begin{smallmatrix} 2 & -3 \\ 3 & 2 \end{smallmatrix}\right)$. La courbe a pour équation dans cette base :

$$13X^2 + 39Y^2 - \frac{4}{\sqrt{13}}(2X - 3Y) + \frac{6}{\sqrt{13}}(3X + 2Y) = 25,$$

qui s'écrit

$$13(X + \frac{1}{\sqrt{13}})^2 + 39Y^2 = 26$$

ou encore

$$(X - X_\Omega)^2 + 3(Y - Y_\Omega)^2 = 2 \tag{6.111}$$

en posant $X_\Omega = -\frac{1}{\sqrt{13}}$ et $Y_\Omega = 0$. On a alors $x_\Omega = -\frac{2}{13} \approx -0.15, y_\Omega = -\frac{3}{13} \approx -0.23$. L'équation (6.111) est l'équation dans la base $(\mathbf{f_1}, \mathbf{f_2})$ d'une ellipse de centre $\Omega$ de coordonnées $(X_\Omega, Y_\Omega)$, grand axe $a = \sqrt{2}$, petit axe $b = \sqrt{\frac{2}{3}} \approx 0.82$. La distance focale c est donnée par $\mathrm{c}^2 = a^2 - b^2 = \frac{4}{3}$ donc $\mathrm{c} = \frac{2}{\sqrt{3}} \approx 1.15$. L'excentricité est $e = \sqrt{\frac{2}{3}} = b$. L'un des foyers $F$ est défini par $\mathbf{\Omega F} = \mathrm{c}\mathbf{f_1} = \frac{2}{\sqrt{3}\sqrt{13}}\left(\begin{smallmatrix} 2 \\ 3 \end{smallmatrix}\right)$, d'où

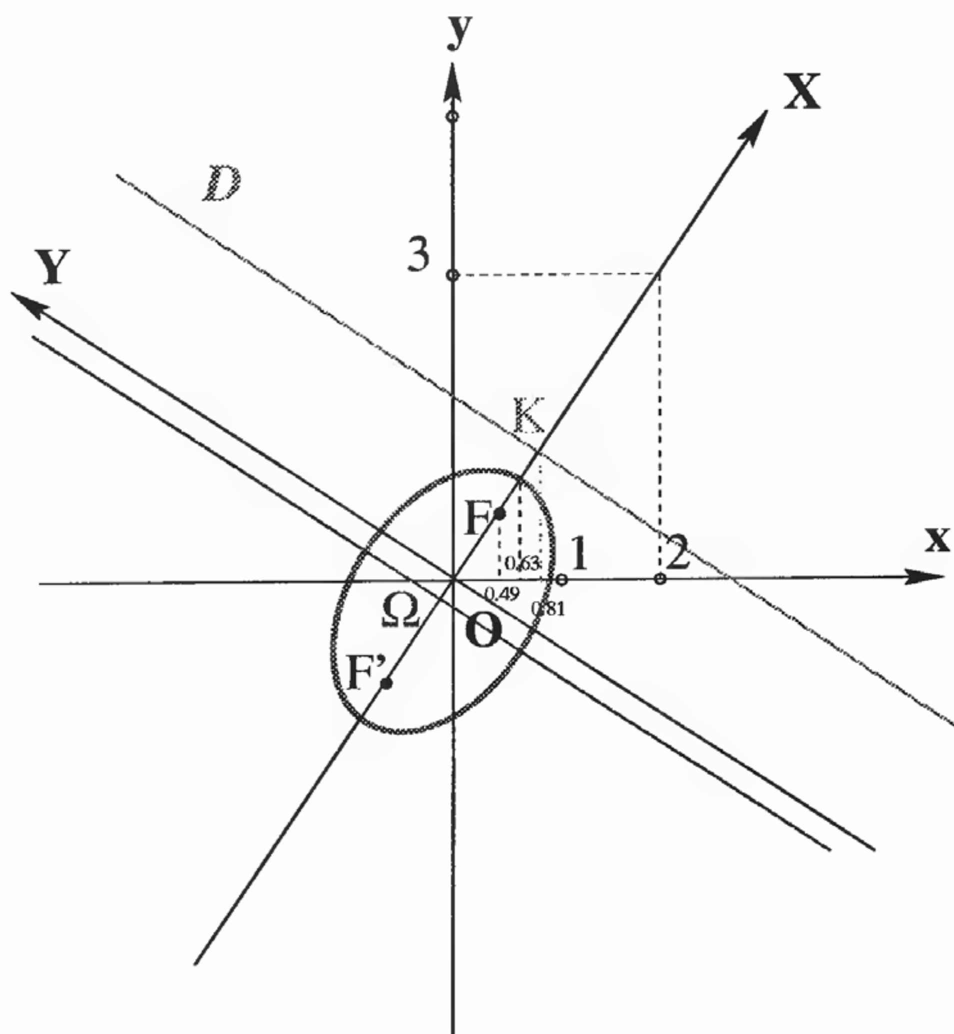



FIG. 6.20: *Exercice 6.5*

$x_F = \frac{2(2\sqrt{13}-\sqrt{3})}{13\sqrt{3}} \approx 0.49$ et $y_F = \frac{3(2\sqrt{13}-\sqrt{3})}{13\sqrt{3}} \approx 0.73$. Le point $K$ d'intersection avec l'axe focal de la directrice $\mathcal{D}$ relative au foyer $F$ est donné par $\mathbf{\Omega K} = \frac{a}{e}\,\mathbf{f_1} = \frac{\sqrt{3}}{\sqrt{13}}\binom{2}{3}$, d'où $x_K = \frac{2(\sqrt{39}-1)}{13} \approx 0.81$ et $y_K = \frac{3(\sqrt{39}-1)}{13} \approx 1.21$. La directrice $\mathcal{D}$ a pour vecteur directeur $\mathbf{f_2}$ ou encore $\binom{-3}{2}$ et passe par le point $K$. Elle a donc pour équation en $x, y$ :

$$\begin{vmatrix} x - \frac{2(\sqrt{39}-1)}{13} & -3 \\ y - \frac{3(\sqrt{39}-1)}{13} & 2 \end{vmatrix} = 0$$

qui s'écrit

$$2x + 3y - (\sqrt{39} - 1) = 0.$$

**Exercice 6.6.**

Quelle est la nature de la courbe d'équation

$$6x^2 + 4xy + 3y^2 - 28x - 14y + 15 = 0\ ?$$

Préciser les coordonnées $(x_\Omega, y_\Omega)$ de son centre $\Omega$, l'excentricité, un foyer et la directrice associée.

**Exercice 6.7.**

On considère pour $\lambda \in \mathbb{R}$ la courbe $\mathcal{C}_\lambda$ d'équation

$$x^2 + 2\lambda xy + 4y^2 - 2x + 2y = 0.$$

(i) Quelle est la nature de la courbe $\mathcal{C}_0$ ? La construire.

(ii) Quelle est la nature de la courbe $\mathcal{C}_2$ ? La construire.

(iii) Montrer que les courbes $\mathcal{C}_\lambda$ passent par 3 points fixes.

(iv) Discuter suivant les valeurs de $\lambda$ la nature de la courbe $\mathcal{C}_\lambda$.

**Exercice 6.8.**

Soit $\mathcal{C}$ une conique. Montrer que l'ensemble des points du plan d'où l'on peut mener deux tangentes perpendiculaires entre elles à $\mathcal{C}$ (lieu *orthoptique de* $\mathcal{C}$) est :

(i) un cercle si $\mathcal{C}$ est une ellipse, et un cercle ou l'ensemble vide si $\mathcal{C}$ est une hyperbole (dans le cas de l'hyperbole, on considère une asymptote comme une *tangente à l'infini* à l'hyperbole);

(ii) la directrice si $\mathcal{C}$ est une parabole.

**Indication.**

(i) Soit $\varepsilon = \pm 1$ et $\mathcal{C}$ (ellipse ou hyperbole suivant la valeur de $\varepsilon$) d'équation

$$\frac{x^2}{a^2} + \varepsilon \frac{y^2}{b^2} = 1,$$

avec $a, b > 0$ , $a^2 - \varepsilon b^2 = c^2$, $c > 0$ étant la distance focale (dans le cas d'une ellipse dégénérée en cercle, on prend $c = 0$). Soit $M_0$ un point du plan de coordonnées $(x_0, y_0)$. Une droite $D$ non verticale, de pente $m \in \mathbb{R}$ passant par $M_0$ a pour vecteur directeur $\mathbf{v} = \begin{pmatrix} 1 \\ m \end{pmatrix}$, et

$$D = \{M_0 + t\mathbf{v};\ t \in \mathbb{R}\}.$$

Les points d'intersection de $D$ et $\mathcal{C}$ sont donnés par l'équation en $t$ :

$$\frac{(x_0 + t)^2}{a^2} + \varepsilon \frac{(y_0 + mt)^2}{b^2} = 1$$

qui s'écrit

$$t^2 \left( \frac{1}{a^2} + \varepsilon \frac{m^2}{b^2} \right) + 2t \left( \frac{x_0}{a^2} + \varepsilon \frac{my_0}{b^2} \right) + \frac{x_0^2}{a^2} + \varepsilon \frac{y_0^2}{b^2} - 1 = 0. \tag{6.112}$$

La droite $D$ est tangente à $\mathcal{C}$ si et seulement si (6.112) est de degré 2 et a une racine double en $t$, *i.e.* les 3 nombres réels $x_0, y_0, m$ vérifient le système des 2 équations suivantes :

$$\frac{1}{a^2} + \varepsilon \frac{m^2}{b^2} \neq 0 \tag{6.113}$$

$$\left( \frac{x_0}{a^2} + \varepsilon \frac{my_0}{b^2} \right)^2 - \left( \frac{1}{a^2} + \varepsilon \frac{m^2}{b^2} \right) \left( \frac{x_0^2}{a^2} + \varepsilon \frac{y_0^2}{b^2} - 1 \right) = 0. \tag{6.114}$$

L'équation (6.114) s'écrit

$$2\varepsilon \frac{mx_0y_0}{a^2b^2} - \varepsilon \frac{y_0^2}{a^2b^2} - \varepsilon \frac{m^2x_0^2}{a^2b^2} + \frac{1}{a^2} + \varepsilon \frac{m^2}{b^2} = 0$$

ou encore

$$m^2(a^2 - x_0^2) + 2mx_0y_0 - y_0^2 + \varepsilon b^2 = 0. \tag{6.115}$$

• Considérons l'équation (6.113). Elle est toujours vérifiée dans le cas $\varepsilon = 1$ (cas de l'ellipse). Mais dans le cas $\varepsilon = -1$ (cas de l'hyperbole), elle s'écrit $m \neq \pm\frac{b}{a}$ et n'est donc pas vérifiée si $m$ est la pente d'une asymptote.

• Considérons maintenant l'équation (6.115) à l'inconnue $m$. Elle est de degré 2 si $x_0 \neq \pm a$. Son discriminant réduit est

$$x_0^2y_0^2 - (a^2 - x_0^2)(-y_0^2 + \varepsilon b^2) = \varepsilon a^2 b^2 \left(\frac{x_0^2}{a^2} + \varepsilon\frac{y_0^2}{b^2} - 1\right).$$

L'équation a 2 solutions réelles distinctes $m_1$ et $m_2$ si et seulement si

$$\varepsilon\left(\frac{x_0^2}{a^2} + \varepsilon\frac{y_0^2}{b^2} - 1\right) > 0. \tag{6.116}$$

La condition (6.116) signifie que $M_0$ appartient à l'extérieur de l'ellipse dans le cas $\varepsilon = 1$. Dans le cas $\varepsilon = -1$, elle signifie que $M_0$ est dans la partie du plan située entre les deux branches de l'hyperbole et contenant le centre $O$.

Dans le cas $\varepsilon = -1$, si $m$ est une solution de (6.115), la droite de pente $m$ passant par $M_0$ n'est pas nécessairement une tangente à l'hyperbole. C'en est une si et seulement si l'équation (6.113) est vérifiée, *i.e.* si $m$ n'est pas la pente d'une asymptote. Si $m$ est la pente d'une asymptote, $m^2 = \frac{b^2}{a^2}$ et l'équation (6.114) s'écrit $\frac{x_0^2}{a^2} - m\frac{y_0^2}{b^2} = 0$, *i.e.* $y_0 = mx_0$. Donc dans ce cas, la droite en question est l'asymptote de pente $m$. Si l'on considère une asymptote comme une *tangente à l'infini* à l'hyperbole, alors pour toute solution $m$ de (6.115), la droite de pente $m$ passant par $M_0$ est une tangente à l'hyperbole. Avec cette convention, il n'y a donc plus lieu de prendre en compte l'équation (6.113).

Avec la convention, pour $x_0 \neq \pm a$, il existe ainsi 2 tangentes (non verticales) à la conique $\mathcal{C}$ issues de $M_0$ perpendiculaires entre elles si et seulement si il existe 2 solutions $m_1, m_2$ de (6.115) telles que les droites passant par $M_0$ et de pentes $m_1, m_2$ soient perpendiculaires, *i.e.*

$$m_1m_2 = -1. \tag{6.117}$$

D'après les relations coefficients-racines, le produit des racines de l'équation (6.115) est $\frac{-y_0^2+\varepsilon b^2}{a^2-x_0^2}$.

Il existe donc 2 racines $m_1$ et $m_2$ de produit $-1$ si et seulement si

$$\frac{-y_0^2 + \varepsilon b^2}{a^2 - x_0^2} = -1 \tag{6.118}$$

*i.e.*

$$x_0^2 + y_0^2 = a^2 + \varepsilon b^2. \tag{6.119}$$

On notera que les deux racines sont alors réelles et distinctes. En effet l'équation (6.115) étant à coefficients réels, elle a soit deux racines réelles distinctes, soit une racine double réelle, soit deux racines complexes conjuguées. Dans les deux derniers cas, la relation (6.117) n'est évidemment pas vérifiée.

On peut aussi voir que les deux racines sont réelles et distinctes en observant que si (6.118) est vérifiée, on a

$$\frac{x_0^2}{a^2} = -\frac{y_0^2}{a^2} + \varepsilon\frac{b^2}{a^2} + 1$$

donc

$$\begin{aligned}\varepsilon\left(\frac{x_0^2}{a^2} + \varepsilon\frac{y_0^2}{b^2} - 1\right) &= \varepsilon\left(y_0^2\left(-\frac{1}{a^2} + \frac{\varepsilon}{b^2}\right) + \varepsilon\frac{b^2}{a^2}\right) \\ &= y_0^2\left(\frac{a^2 - \varepsilon b^2}{a^2b^2}\right) + \frac{b^2}{a^2} \\ &= y_0^2\frac{c^2}{a^2b^2} + \frac{b^2}{a^2} > 0\,.\end{aligned}$$

et (6.116) est vérifiée.

L'ensemble des points $M_0$ du plan de coordonnées $(x_0, y_0)$ tels que $x_0 \neq \pm a$ et d'où sont issues 2 tangentes non verticales perpendiculaires entre elles est donc l'ensemble défini par (6.119). On va maintenant distinguer les deux cas ellipse-hyperbole.

► 1er cas : $\varepsilon = 1$ (cas de l'ellipse).

$$\{M_0;\ x_0^2 + y_0^2 = a^2 + b^2,\ x_0 \neq \pm a\}$$

est le cercle de centre l'origine $O$ et de rayon $\sqrt{a^2+b^2}$ privé des 4 points $x_0 = \pm a, y_0 = \pm b$. Mais pour $x_0 = \pm a$, la droite verticale $x = x_0$ est tangente à l'ellipse, et il passe une seconde tangente horizontale par $M_0$ si et seulement si $y_0 = \pm b$. Les 4 points exclus sont donc les points d'où l'on peut mener deux tangentes perpendiculaires dont l'une est verticale.

L'ensemble des points d'où l'on peut mener deux tangentes perpendiculaires entre elles à l'ellipse est donc le cercle entier

$$\{M_0;\ x_0^2 + y_0^2 = a^2 + b^2\}.$$

► 2ème cas : $\varepsilon = -1$ (cas de l'hyperbole).

$$\{M_0;\ x_0^2 + y_0^2 = a^2 - b^2,\ x_0 \neq \pm a\} = \{M_0;\ x_0^2 + y_0^2 = a^2 - b^2\}$$

est l'ensemble vide si $a < b$, le cercle réduit à l'origine $\{O\}$ si $a = b$, et le cercle de rayon $\sqrt{a^2 - b^2}$ si $a > b$. Par ailleurs, si $x_0 = \pm a$, la verticale $x = x_0$ est une tangente à l'hyperbole, mais l'hyperbole n'ayant pas de tangente horizontale, le point $M_0$ ne convient pas.

L'ensemble des points d'où l'on peut mener deux tangentes perpendiculaires entre elles à l'hyperbole est donc l'ensemble vide si $a < b$, et si $a \geqslant b$, c'est le cercle de rayon $\sqrt{a^2 - b^2}$.

(ii) Soit $\mathcal{P}$ la parabole d'équation

$$y^2 = 2px.$$

Soit $M_0$ un point du plan de coordonnées $(x_0, y_0)$. Une droite $D$ non verticale, de pente $m \in \mathbb{R}$ passant par $M_0$ a pour vecteur directeur $\mathbf{v} = \begin{pmatrix}1\\ m\end{pmatrix}$, et l'on a

$$D = \{M_0 + t\mathbf{v};\ t \in \mathbb{R}\}.$$

Les points d'intersection de $D$ et $C$ sont donnés par l'équation en $t$ :

$$(y_0 + mt)^2 - 2p(x_0 + t) = 0$$

qui s'écrit

$$m^2t^2 + 2t(y_0m - p) + y_0^2 - 2px_0 = 0. \tag{6.120}$$

La droite $D$ est tangente à $\mathcal{P}$ si et seulement si (6.120) est de degré 2 et a une racine double en $t$, *i.e.*

$$m \neq 0 \tag{6.121}$$
$$(y_0m - p)^2 - m^2(y_0^2 - 2px_0) = 0. \tag{6.122}$$

L'équation (6.122) s'écrit après simplification par $p$ :

$$2x_0m^2 + 2y_0m + p = 0 \tag{6.123}$$

Les pentes des tangentes non verticales issues de $M_0$ sont donc les $m \in \mathbb{R}, m \neq 0$ qui vérifient (6.123). Comme aucune horizontale n'est tangente à $\mathcal{P}$, il y a deux tangentes issues de $M_0$ et perpendiculaires entre elles si et seulement si l'équation (6.123) a deux racines réelles $m_1, m_2$ telles que

$$m_1m_2 = -1$$

D'après les relations coefficients-racines, cela équivaut à l'équation

$$\frac{p}{2x_0} = -1$$

*i.e.*

$$x_0 = -\frac{p}{2} \tag{6.124}$$

Or (6.124) est l'équation de la directrice $\mathcal{D}$ de $\mathcal{P}$, d'où le résultat.

**Exercice 6.9.**

Dans le plan affine euclidien rapporté au repère orthonormé $(O, (\mathbf{e_1}, \mathbf{e_2}))$, soient les 2 points $A(-\sqrt{2}, -\sqrt{2})$, $B(\sqrt{2}, -\sqrt{2})$ et $\Gamma$ le cercle de centre $O$ et de rayon 2. Soit $M$ un point de $\Gamma$ différent de $A$ et $B$. On note $P$ et $Q$ les points d'intersection avec l'axe des abscisses des droites $MA$ et $MB$ respectivement, et $I$ le centre du cercle $\Gamma_M$ circonscrit au triangle $MPQ$. Soit enfin $\theta$ la détermination principale de l'angle orienté $(\widehat{\mathbf{e_1}, \mathbf{OM}})$.

(i) Montrer que le triangle $MPQ$ se déduit du triangle $MAB$ par l'homothétie de centre $M$ et de rapport $\lambda_M = \frac{\sqrt{2}\sin\theta}{1+\sqrt{2}\sin\theta}$, et en déduire que le cercle $\Gamma_M$ se déduit du cercle $\Gamma$ par cette homothétie.

(ii) Calculer $\mathbf{OI}$ en fonction de $\theta$ et donner l'équation en coordonnées polaires du lieu $\mathcal{H}$ du point $I$ lorsque $M$ décrit $\Gamma$.

(iii) Quelle est la nature de $\mathcal{H}$?

**Indication.**

Voir Fig. 6.21.

(i) La droite $AB$ étant parallèle à l'axe des $x$, le triangle $MPQ$ se déduit du triangle $MAB$ par une homothétie de centre $M$ et de rapport $\lambda_M$. On a

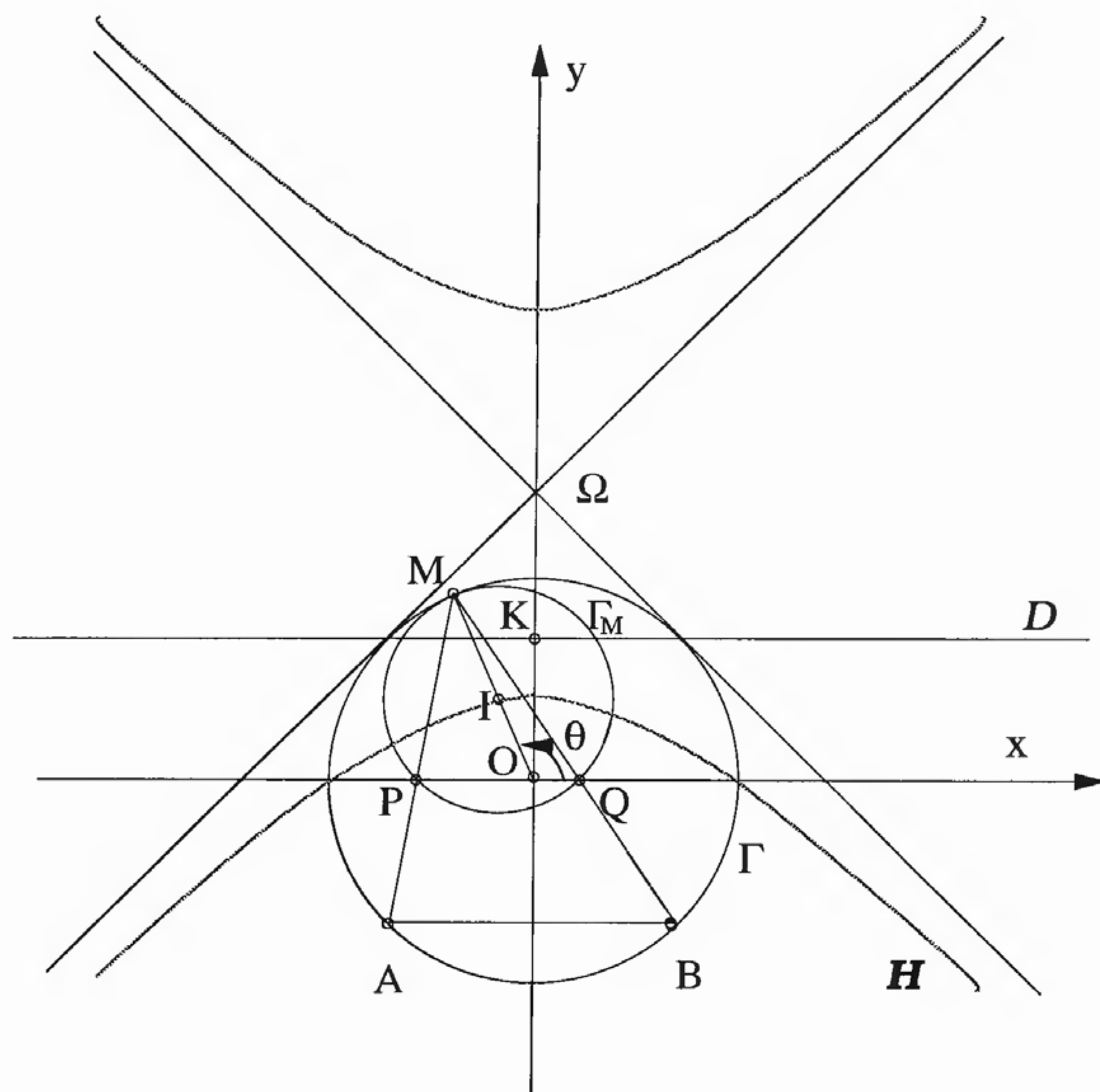



FIG. 6.21: *Exercice 6.9*

$\mathbf{OM} = \binom{2\cos\theta}{2\sin\theta}$, $\mathbf{MA} = \binom{-\sqrt{2}-2\cos\theta}{-\sqrt{2}-2\sin\theta}$, et si $x_P$ désigne l'abscisse de $P$, $\mathbf{MP} = \binom{x_P-2\cos\theta}{-2\sin\theta}$, donc l'équation $\mathbf{MP} = \lambda_M\mathbf{MA}$ donne par projection sur la deuxième composante $-2\sin\theta = -\lambda_M(\sqrt{2}+2\sin\theta)$ *i.e.* $\lambda_M = \frac{\sqrt{2}\sin\theta}{1+\sqrt{2}\sin\theta}$. Par cette homothétie, le cercle $\Gamma$ passant par $A, B, M$ est transformé en un cercle passant par $P, Q, M$, *i.e.* en $\Gamma_M$. Le centre $I$ de $\Gamma_M$ est donc le transformé du centre $O$ du cercle $\Gamma$.

(ii) $\mathbf{OI} = \mathbf{OM} + \mathbf{MI} = (1-\lambda_M)\mathbf{OM} = \frac{1}{1+\sqrt{2}\sin\theta}\mathbf{OM} = \frac{2}{1+\sqrt{2}\sin\theta}\mathbf{u}(\theta)$ avec $\mathbf{u}(\theta) = \cos\theta\,\mathbf{e_1} + \sin\theta\,\mathbf{e_2}$. L'équation en coordonnées polaires de $\mathcal{H}$ est donc

$$r = \frac{2}{1+\sqrt{2}\sin\theta}. \tag{6.125}$$

(iii) L'équation (6.125) s'écrit encore avec $\theta_0 = \frac{\pi}{2}$ :

$$r = \frac{2}{1+\sqrt{2}\cos(\theta-\theta_0)}.$$

C'est l'équation d'une hyperbole d'excentricité $e = \sqrt{2}$, de foyer $O$, d'axe focal (orienté vers la directrice $\mathcal{D}$ relative au foyer $O$) $Oy$, de paramètre $p = 2$. Comme $p = ed$ où $d$ est la distance du foyer à la directrice associée, $d = \sqrt{2}$ et la directrice $\mathcal{D}$ a pour équation $y = \sqrt{2}$. Par ailleurs, si $a$ est le grand axe et $c = a\sqrt{2}$ la distance focale, $a = c - r(\frac{\pi}{2}) = c - \frac{2}{1+\sqrt{2}}$ donc $a = 2$ et $c = 2\sqrt{2}$. Le centre $\Omega$ de l'hyperbole a pour coordonnées $x_\Omega = 0$, $y_\Omega = 2\sqrt{2}$. Les asymptotes correspondent à $\sin\theta = -\frac{\sqrt{2}}{2}$, *i.e.* $\theta = -\frac{\pi}{4} \pmod{2\pi}$ et $\theta = \pi - (-\frac{\pi}{4}) = \frac{5\pi}{4} \pmod{2\pi}$. Ce sont les angles polaires des 2 points $A$ et $B$. On vérifie facilement que les asymptotes sont les deux tangentes au cercle $\Gamma$ issues du point $\Omega$. En effet, si une droite passant par $\Omega$ est tangente à $\Gamma$ en un point $T$, dans le triangle rectangle $\Omega TO$, on a $\|\mathbf{OT}\| = \|\mathbf{O\Omega}\|\sin\alpha$ avec $\alpha$ l'angle non orienté des deux vecteurs $\mathbf{O\Omega}$ et $\mathbf{OT}$. Cette relation donne $\sin\alpha = \frac{\sqrt{2}}{2}$, donc $\alpha = \frac{\pi}{4}$ (puisque $\alpha < \frac{\pi}{2}$). La directrice $\mathcal{D}$ passe par les points de tangence car c'est la médiatrice du segment $[O\Omega]$. L'hyperbole $\mathcal{H}$ est équilatère (*i.e.* les asymptotes sont orthogonales). Le point $I$ décrit la branche de $\mathcal{H}$ qui rencontre $\Gamma$ lorsque $M$ décrit l'arc de $\Gamma$ situé "au dessus de $A, B$" et l'autre branche lorsque $M$ décrit l'arc de $\Gamma$ situé "au dessous de $A, B$".

**Exercice 6.10.**

Tracer la courbe dont l'équation en coordonnées polaires est

$$r = \frac{1}{1+2\cos\theta}$$

**Indication.**

D'après le Th.6.9 , la courbe est une hyperbole d'excentricité 2, de paramètre 1, le pôle $O$ est l'un des foyers, l'axe polaire est l'axe focal, orienté vers la directrice associée au foyer $O$. D'après la relation $p = ed$, la directrice est la verticale au point d'abscisse $x = \frac{1}{2}$.

Le domaine de définition de la fonction $\theta \mapsto r(\theta) = \frac{1}{1+2\cos\theta}$ est $\mathbb{R} \setminus \{\pm\frac{2\pi}{3} + 2k\pi\,;\, k \in \mathbb{Z}\}$. La fonction est $2\pi$-périodique, donc on l'étudie sur un intervalle de longueur $2\pi$, par exemple $[-\pi, \pi]$. La fonction étant paire, on restreint l'intervalle d'étude à $[0, \pi] \setminus \{\frac{2\pi}{3}\}$ et on complètera par une symétrie par rapport à l'axe des abscisses. On a $r'(\theta) = \frac{2\sin\theta}{(1+2\cos\theta)^2}$, d'où le tableau de variations.

| θ | 0 | | π/2 | | 2π/3 | | π |
|---|---|---|---|---|---|---|---|
| r' | 0 | + | 2 | + | ‖ | + | 0 |
| r | 1/3 | ↗ | 1 | ↗ | +∞ ‖ −∞ | ↗ | −1 |
| r/r' | +∞ | | 1/2 | | ‖ | | −∞ |

Il y a une branche infinie de direction $\theta = \frac{2\pi}{3}$. Soit $X, Y$ les coordonnées associées au *repère mobile*

$$\mathcal{R}_{\frac{2\pi}{3}} = \left(\mathbf{u}_{\frac{2\pi}{3}} = \cos\frac{2\pi}{3}\,\mathbf{e_1} + \sin\frac{2\pi}{3}\,\mathbf{e_2},\ \mathbf{w}_{\frac{2\pi}{3}} = -\sin\frac{2\pi}{3}\,\mathbf{e_1} + \cos\frac{2\pi}{3}\,\mathbf{e_2}\right),$$

$\mathbf{e_1}$ étant un vecteur unitaire directeur de l'axe polaire orienté et $\mathbf{e_2}$ le vecteur directement perpendiculaire. On sait que si $r\,\sin(\theta - \frac{2\pi}{3})$ possède une limite $Y_0$ quand $\theta \to \frac{2\pi}{3}$, alors la droite ayant pour équation $Y = Y_0$ dans le repère mobile $(\mathbf{u}_{\frac{2\pi}{3}}, \mathbf{w}_{\frac{2\pi}{3}})$ est asymptote à la courbe. La position par rapport à l'asymptote est donnée dans le repère mobile par le signe de $r\,\sin(\theta - \frac{2\pi}{3}) - Y_0$. On a ici en posant $\theta - \frac{2\pi}{3} = h$

$$\begin{aligned} r\,\sin h &= \frac{\sin h}{1 + 2\cos\left(h + \frac{2\pi}{3}\right)} \\ &= \frac{\sin h}{1 - \cos h - \sqrt{3}\,\sin h} \\ &= \frac{h + o(h^2)}{-h\sqrt{3} + \frac{h^2}{2} + o(h^2)} \\ &= -\frac{1}{\sqrt{3}} - \frac{h}{6} + o(h)\,. \end{aligned}$$

Donc il y a une asymptote d'équation $Y = -\frac{1}{\sqrt{3}}$. Notons $A$ le point de coordonnées $(0, -\frac{1}{\sqrt{3}})$ dans le repère mobile $\mathcal{R}_{\frac{2\pi}{3}}$. La position par rapport à l'asymptote est donnée dans le repère mobile par le signe du terme $-\frac{h}{6}$. Quand $h \to 0$ par valeur négatives, la courbe est au dessus, quand $h \to 0$ par valeur positives, la courbe est au dessous. Soit $\Omega$ le centre de l'hyperbole, intersection des asymptotes. Comme le triangle $O\Omega A$ est rectangle en $A$, l'abscisse de $\Omega$ est donné par $x_\Omega \cos\frac{\pi}{6} = \frac{1}{\sqrt{3}}$, soit $x_\Omega = \frac{2}{3}$. D'où le tracé de l'hyperbole (voir Fig. 6.22).

**Exercice 6.11.**

Le plan affine euclidien est rapporté au repère orthonormé $(O, (\mathbf{e_1}, \mathbf{e_2}))$. Soit $\mathcal{C}$ un cercle d'équation

$$x^2 + y^2 - 2ax - 2by + c = 0.$$

passant par le point $A$ de coordonnées $(x_0, -1)$.

(i) Montrer que la pente de la tangente en $A$ à $\mathcal{C}$ est :

$$\frac{x_0 - a}{b + 1}.$$

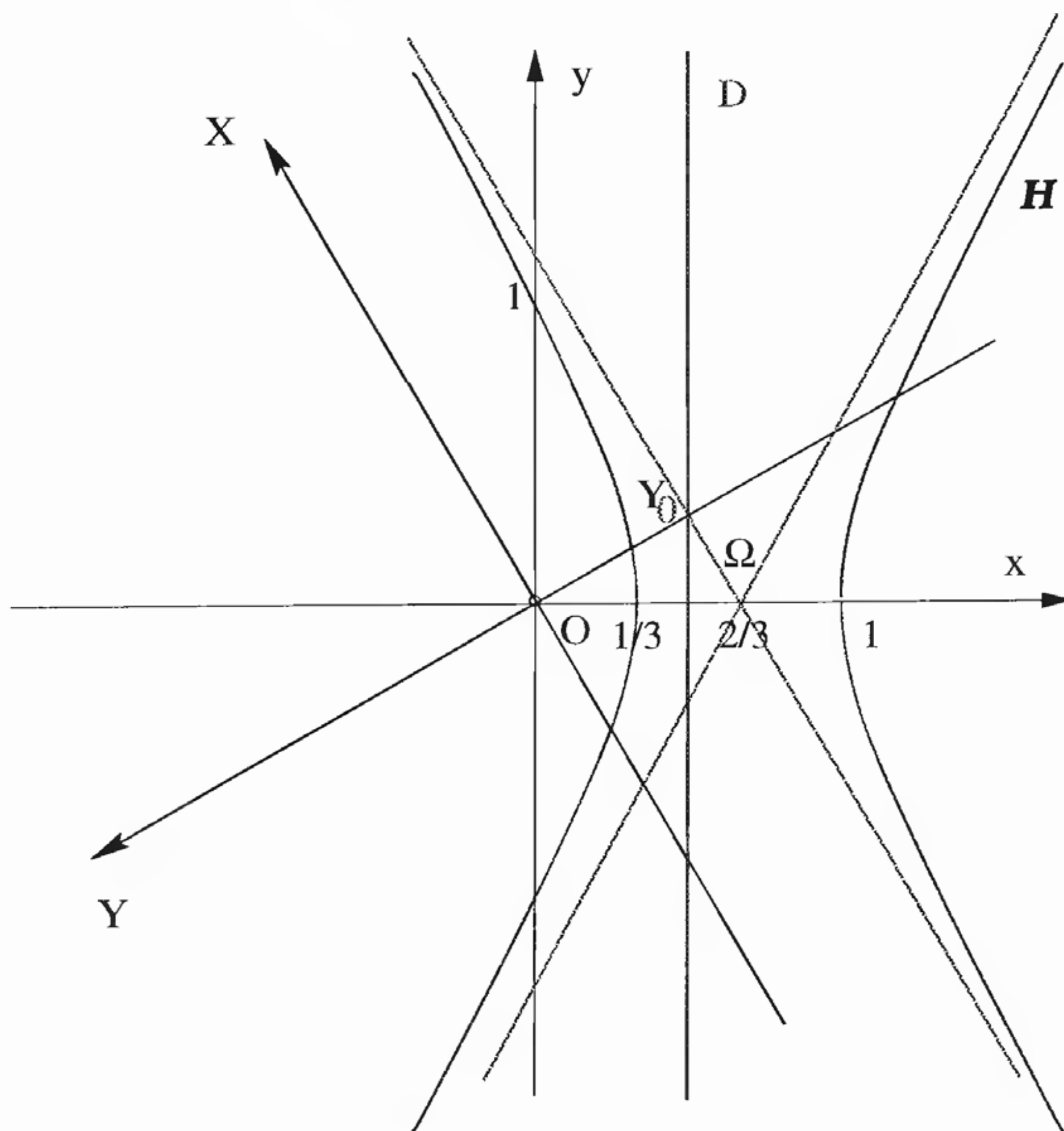



FIG. 6.22: *Exercice 6.10*

(ii) Montrer que le lieu des centres des cercles $\mathcal{C}$ passant par le point $F$ de coordonnées $(0,2)$, ayant un point d'intersection avec la droite d'équation $y = -1$ tel que la tangente en ce point ait une pente égale à $\sqrt{3}$, est la courbe $\mathcal{H}$ d'équation

$$x^2 - 3y^2 - 12y = 0.$$

(iii) Quelle est la nature de $\mathcal{H}$? Préciser un foyer et la directrice associée. Dessiner $\mathcal{H}$.

**Exercice 6.12.**

Dans le plan affine euclidien rapporté au repère orthonormé $(O, (\mathbf{e_1}, \mathbf{e_2}))$, soit $A$ le point de coordonnées $(a, 0)$, $a \in \mathbb{R}$. Soit $f : I \to \mathbb{R}$ une fonction de classe $C^1$ sur un intervalle $I$ de $\mathbb{R}$ et $\mathcal{C}$ la courbe d'équation $y = f(x)$, $x \in I$. Soit $x \in I$, $M$ le point de coordonnées $(x, f(x))$, et $T$ le point d'intersection (supposé exister) de la tangente en $M$ à la courbe $\mathcal{C}$ avec l'axe des ordonnées.

(i) Montrer que les droites $AM$ et $AT$ sont orthogonales au point $A$ si et seulement si $f$ vérifie

$$xf(x)f'(x) - (f(x))^2 + ax - a^2 = 0.$$

(ii) En posant $z = y^2$, intégrer l'équation différentielle

$$xyy' - y^2 + ax - a^2 = 0.$$

et en déduire que les droites $AM$ et $AT$ sont orthogonales au point $A$ quel que soit $x \in I$ si et seulement si il existe $C \in \mathbb{R}$ tel que

$$(f(x))^2 = Cx^2 + 2ax - a^2 \ \forall x \in I.$$

(iii) Soit $\Gamma_C$ la courbe d'équation cartésienne

$$y^2 = Cx^2 + 2ax - a^2.$$

Montrer que $\Gamma_C$ est une conique dont $A$ est un foyer. Quelle est la directrice associée? Exprimer l'excentricité en fonction de $C$.

**Exercice 6.13.**

Soient $A$ et $B$ deux points du plan affine euclidien et $k > 0, k \neq 1$. Montrer que l'ensemble des points $M$ tels que $\frac{\|\mathbf{MA}\|}{\|\mathbf{MB}\|} = k$ est le cercle ayant pour centre le barycentre $G$ des points $A$ et $B$ affectés respectivement des poids 1 et $-k^2$, et pour rayon $R = \frac{k}{|1-k^2|}\|\mathbf{AB}\|$.

**Indication.**

Utilisons un repère orthonormé $(O, (\mathbf{e_1}, \mathbf{e_2}))$, avec $O$ le milieu du segment $[A, B]$ et $\mathbf{e_1} = \frac{\mathbf{OB}}{\|\mathbf{OB}\|}$. Les coordonnées des points $A$ et $B$ sont respectivement $(-\alpha, 0)$ et $(\alpha, 0)$ $(\alpha > 0)$. Soit $M$ de coordonnées $(x, y)$. L'équation

$$\frac{\|\mathbf{MA}\|}{\|\mathbf{MB}\|} = k \qquad (6.126)$$

s'écrit

$$\begin{aligned} x^2 + \alpha^2 + 2\alpha x + y^2 &= k^2(x^2 + \alpha^2 - 2\alpha x + y^2) \\ &\Leftrightarrow \\ x^2(1-k^2) + 2\alpha x(1+k^2) + y^2(1-k^2) &= (k^2-1)\alpha^2 \\ &\Leftrightarrow \\ \left(x + \alpha\frac{1+k^2}{1-k^2}\right)^2 + y^2 &= \alpha^2\left(\frac{1+k^2}{1-k^2}\right)^2 - \alpha^2 \\ &\Leftrightarrow \\ \left(x + \alpha\frac{1+k^2}{1-k^2}\right)^2 + y^2 &= \frac{4k^2\alpha^2}{(1-k^2)^2}. \end{aligned}$$

On reconnaît l'équation d'un cercle dont le centre $G$ a pour coordonnées

$$(-\alpha\frac{1+k^2}{1-k^2}, 0)$$

*i.e.*

$$\mathbf{OG} = \begin{pmatrix} -\alpha\frac{1+k^2}{1-k^2} \\ 0 \end{pmatrix}.$$

Le rayon $R$ est

$$R = \frac{2k\alpha}{|1-k^2|} = \frac{k}{|1-k^2|}\|\mathbf{AB}\|.$$

On a

$$\mathbf{OG} = \frac{1}{1-k^2}\left(\mathbf{OA} - k^2\mathbf{OB}\right)$$

donc $G$ est le barycentre des points $A$ et $B$ affectés respectivement des poids 1 et $-k^2$. Quand $k$ varie sur $[0,1[$, l'abscisse $x_G = -\alpha\frac{1+k^2}{1-k^2}$ de $G$ varie de $-\alpha$ à $-\infty$ donc $G$ décrit la demi-droite orientée $[A, -\infty[$ de l'axe des $x$. Quand $k$ varie sur $]1, +\infty[$, $x_G$ varie de $+\infty$ à $\alpha$ donc $G$ décrit la demi-droite orientée $]+\infty, B[$ de l'axe des $x$. La valeur $k = 0$ donne le cercle-point $\{A\}$. Le point $B$ est un point limite correspondant à $k \to +\infty$. On a

$$R^2 = \frac{4\alpha^2k^2}{(1-k^2)^2} = \alpha^2\frac{(1+k^2)^2 - (1-k^2)^2}{(1-k^2)^2} = \alpha^2\left(\frac{1+k^2}{1-k^2}\right)^2 - \alpha^2 = \|\mathbf{OG}\|^2 - \alpha^2$$

donc

$$\|\mathbf{OG}\|^2 = R^2 + \alpha^2.$$

Cela signifie que le cercle $\Gamma$ correspondant à la valeur $k$ est orthogonal au cercle de centre $O$ et de rayon $\alpha$. Or si l'on considère les deux cercle-points $\Gamma_A = \{A\}$ et $\Gamma_B = \{B\}$, l'équation (6.126) s'écrit

$$\begin{aligned}
\mathcal{P}_{\Gamma_A}(M) &= k^2\mathcal{P}_{\Gamma_B}(M) \\
&\Leftrightarrow \\
\mathcal{P}_{\Gamma_A}(M) - k^2\mathcal{P}_{\Gamma_B}(M) &= 0 \\
&\Leftrightarrow \\
\frac{1}{1-k^2}\mathcal{P}_{\Gamma_A}(M) - \frac{k^2}{1-k^2}\mathcal{P}_{\Gamma_B}(M) &= 0 \\
&\Leftrightarrow \\
\lambda\mathcal{P}_{\Gamma_A}(M) + (1-\lambda)\mathcal{P}_{\Gamma_B}(M) &= 0 \quad (\lambda = \frac{1}{1-k^2} \neq 0).
\end{aligned}$$

On voit donc que les cercles en question forment quand $k$ varie le faisceau de cercles défini par $\Gamma_A = \{A\}$ et $\Gamma_B = \{B\}$, *i.e.* le faisceau à points limites $A$ et $B$, privé du cercle-point $\Gamma_B = \{B\}$ car $\lambda \neq 0$. On sait que ce faisceau est l'ensemble des cercles dont le centre est sur l'axe des $x$ (et ici $\neq B$) et qui sont orthogonaux au cercle de centre $O$ et de rayon $\alpha$. L'axe radical du faisceau est la médiatrice du segment $[A, B]$ (axe des $y$).

**Exercice 6.14.**

Un astéroïde a une orbite autour du Soleil coplanaire avec l'orbite de la Terre, l'orbite de la Terre étant supposée *circulaire* (on néglige l'interaction astéroïde-Terre). On suppose que l'astéroïde passe à une distance minimale du Soleil égale à la moitié du rayon de l'orbite de la Terre, sa vitesse numérique étant alors le double de celle de la Terre.

(i) L'astéroïde quittera-t'il le système solaire?

(ii) Calculer la vitesse numérique de l'astéroïde à l'endroit où il croisera l'orbite terrestre, ainsi que l'angle entre les deux orbites.

**Indication.**

Voir Fig. 6.23.

(i) Soient $S$ le Soleil, $T$ la Terre, $M$ l'astéroïde, $\mu$ la masse du Soleil, $m$ celle de la Terre, $m'$ celle de l'astéroïde. Le champ de gravitation créé par le Soleil en un point $P$ est $G_P = -\frac{\mathcal{G}\mu}{r^2}\mathbf{u}$ avec $\mathbf{u} = \frac{\mathbf{SP}}{\|\mathbf{SP}\|}$ et $r = \|\mathbf{SP}\|$. La force exercée sur la Terre par le

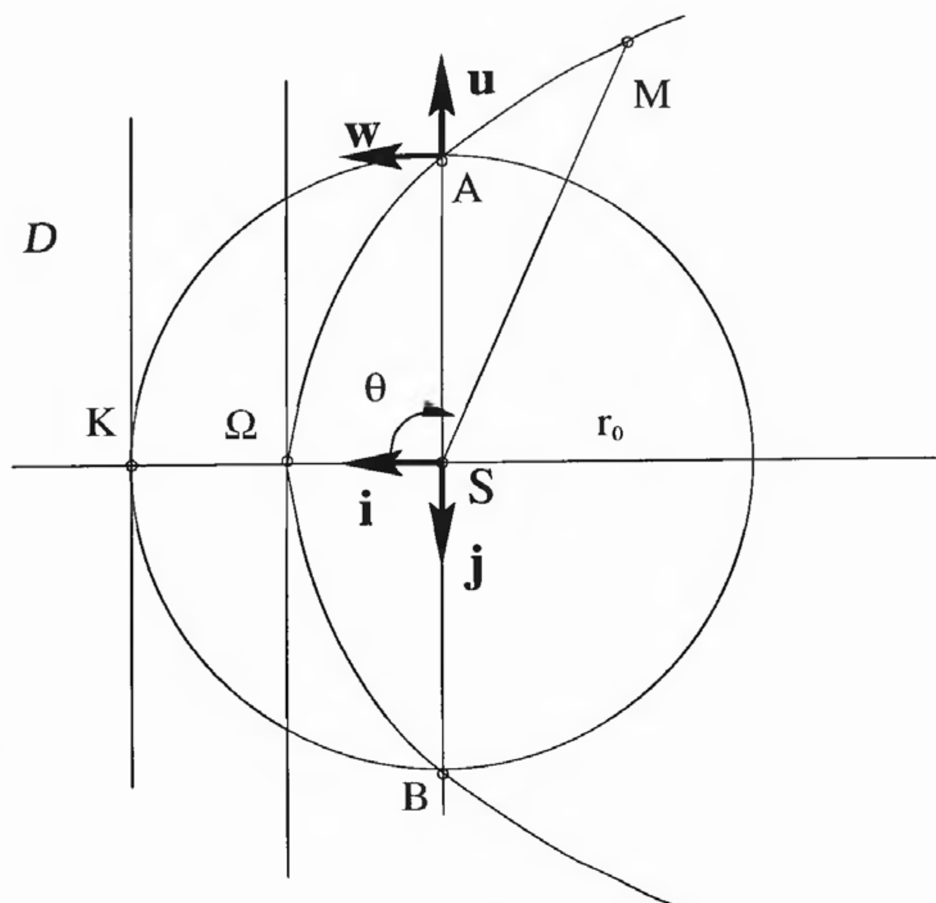



FIG. 6.23: *Exercice 6.14*

Soleil est donc $F_T = -\frac{k}{r^2}\mathbf{u}$ avec $\mathbf{u} = \frac{\mathbf{ST}}{\|\mathbf{ST}\|}$, $r = \|\mathbf{ST}\|$ et $k = \mathcal{G}\mu m$. Comme l'orbite de la Terre est supposée circulaire, la vitesse numérique de la Terre est constante égale à $v_0$, et

$$v_0^2 = \frac{k}{mr_0} = \frac{\mathcal{G}\mu}{r_0}.$$

Soit $\Omega$ (périhélie) le point où l'astéroïde $M$ est le plus près du Soleil. Sa vitesse numérique en $\Omega$ est par hypothèse $v_\Omega = 2v_0$. L'énergie totale de l'astéroïde en $\Omega$ est donc :

$$E = \frac{1}{2}m'v_\Omega^2 - \frac{\mathcal{G}\mu m'}{\frac{r_0}{2}} = 2m'v_0^2 - 2\frac{\mathcal{G}\mu m'}{r_0}.$$

Elle est nulle. Cela signifie que l'astéroïde a une trajectoire parabolique. Il quittera donc le système solaire.

(ii) L'énergie totale de l'astéroïde en $A$ ou $B$, points d'intersection de son orbite avec celle de la Terre, est :

$$E = \frac{1}{2}m'v_A^2 - \frac{\mathcal{G}\mu m'}{r_0}$$

donc comme elle est nulle, $v_A = \sqrt{2}v_0$.

La trajectoire de l'astéroïde a pour équation en coordonnées polaires

$$r = \frac{p}{1+\cos\theta}$$

où $p$ est le paramètre de la parabole. On sait que $\|\Omega\mathbf{S}\| = \frac{p}{2}$, puisque le foyer est $S$ et le sommet $\Omega$. Donc $p = r_0$. La parabole coupe l'orbite de la Terre pour $r = r_0$, *i.e.* $\cos\theta = 0$. Les points $A$ et $B$ ont donc pour angles polaires respectifs $\theta_A = -\frac{\pi}{2}$ et $\theta_B = \frac{\pi}{2} \pmod{2\pi}$.

Soit $t_A$ l'instant où l'astéroïde passe au point $A$. Sa vitesse est donnée dans le *repère mobile* $(\mathbf{u}, \mathbf{w})$ au temps $t_A$ (cf. (6.66)) par la formule (6.69) :

$$\mathbf{V} = \dot{r}\,\mathbf{u} + r\dot{\theta}\,\mathbf{w}.$$

La vitesse de la Terre au point $A$ est parallèle à $\mathbf{w}$. La détermination principale $\psi = \widehat{(\mathbf{u}, \mathbf{V})}$ de l'angle $\widehat{(\mathbf{u}, \mathbf{V})}$ est telle que

$$\tan\psi = \left[\frac{r\dot{\theta}}{\dot{r}}\right]_{t=t_A} = \left[\frac{r}{\frac{dr}{d\theta}}\right]_{\theta=-\frac{\pi}{2}} = \left[\frac{1+\cos\theta}{\sin\theta}\right]_{\theta=-\frac{\pi}{2}} = -1,$$

donc $\psi = -\frac{\pi}{4} \pmod{\pi}$. La détermination principale $\varphi = \widehat{(\mathbf{w}, \mathbf{V})}$ de l'angle $\widehat{(\mathbf{w}, \mathbf{V})}$ est telle que $\varphi \equiv \psi - \frac{\pi}{2} \pmod{2\pi}$, donc $\varphi \equiv \frac{\pi}{4} \pmod{\pi}$. Au point $B$, puisque $\theta_B = \frac{\pi}{2}$, on obtiendrait de même $\psi = \frac{\pi}{4} \pmod{\pi}$ et $\varphi \equiv -\frac{\pi}{4} \pmod{\pi}$.

# Chapitre 7

# Angles en géométrie euclidienne plane.

## 7.1 Angle orienté de 2 vecteurs.

### 7.1.1 Un lemme fondamental.

Soit $E$ un espace vectoriel euclidien de dimension 2. Rappelons qu'un endomorphisme isométrique de déterminant 1 de $E$ est appelé une rotation de $E$ et que le groupe $SO(E)$ des rotations de $E$ est isomorphe au groupe $SO(2)$ (Th. 2.3) et est donc commutatif (Th. 2.5).

**Lemme 7. 1.** *Soient* $\mathbf{u}, \mathbf{v}$ *deux vecteurs unitaires de* $E$. *Il existe* $f \in SO(E)$ *unique tel que* $f(\mathbf{u}) = \mathbf{v}$.

**Démonstration.**

Introduisons une base orthonormée $\mathcal{B} = (\mathbf{e_1}, \mathbf{e_2})$ de $E$. Soit $\binom{x}{y}$ et $\binom{x'}{y'}$ les composantes de $\mathbf{u}$ et $\mathbf{v}$ respectivement dans la base $\mathcal{B}$. Un endomorphisme $f$ de $E$ est un élément de $SO(E)$ si et seulement si sa matrice $A$ dans la base $\mathcal{B}$ est un élément de $SO(2)$. Or on sait que

$$SO(2) = \{\begin{pmatrix} a & -b \\ b & a \end{pmatrix};\ a, b \in \mathbb{R},\ a^2 + b^2 = 1\}.$$

Tout revient donc à démontrer qu'il existe une telle matrice

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix};\ a, b \in \mathbb{R},\ a^2 + b^2 = 1 \tag{7.1}$$

vérifiant $A\binom{x}{y} = \binom{x'}{y'}$. Cela s'écrit :

$$\begin{aligned} ax - by &= x' \\ bx + ay &= y' \end{aligned}$$

ou encore

$$\begin{aligned} xa - yb &= x' \\ ya + xb &= y' \end{aligned}$$

C'est un système d'équations où les inconnues sont $a$ et $b$. Le déterminant du système est $\Delta = x^2 + y^2 = \|\mathbf{u}\|^2 = 1$. Il y a donc une solution unique $a = \begin{vmatrix} x' & -y \\ y' & x \end{vmatrix} = xx' + yy'$, $b = \begin{vmatrix} x & x' \\ y & y' \end{vmatrix} = xy' - yx'$. On a bien $a^2 + b^2 = 1$ puisque

$$\begin{aligned} a^2 + b^2 = (xx' + yy')^2 + (xy' - yx')^2 &= x^2(x'^2 + y'^2) + y^2(x'^2 + y'^2) \\ &= (x^2 + y^2)(x'^2 + y'^2) = \|\mathbf{u}\|^2\|\mathbf{v}\|^2 = 1. \end{aligned}$$

D'où le résultat. □

**Remarque.** Ce lemme ne se généralise pas aux dimensions $n > 2$, car l'unicité est alors perdue. Par exemple, pour $n = 3$, si $\mathbf{u}$ est un vecteur normé et que l'on considère $\mathbf{v} = -\mathbf{u}$, alors pour tout vecteur normé $\mathbf{k}$ orthogonal à $\mathbf{u}$, la rotation $R_{\mathbf{k}}(\pi)$ d'axe orienté $\mathbb{R}\mathbf{k}$ et d'angle $\pi$ vérifie $R_{\mathbf{k}}(\pi)(\mathbf{u}) = \mathbf{v}$.

### 7.1.2 Identification canonique $SO(E) \cong SO(2)$ dans le cas orienté.

On peut toujours identifier $SO(E)$ à $SO(2)$ dès que l'on fixe une base orthonormée de $E$. Mais il y a une infinité de bases orthonormées de $E$ et l'on pourrait donc s'attendre à ce que cela donne une infinité d'identifications différentes. Nous allons voir qu'il n'y en a en fait que 2 différentes, et qu'elles correspondent aux 2 orientations possibles sur $E$.

**Lemme 7. 2.** *Soit $f$ une rotation de $E$, $\mathcal{B}, \mathcal{B}'$ deux bases orthonormées de $E$, et $A, B \in SO(2)$ les matrices de $f$ dans les bases $\mathcal{B}, \mathcal{B}'$ respectivement. Alors :*
- *Si $\mathcal{B}$ et $\mathcal{B}'$ appartiennent à la même orientation de $E$, on a $B = A$;*
- *Si $\mathcal{B}$ et $\mathcal{B}'$ n'appartiennent pas à la même orientation de $E$, on a $B = A^{-1}$.*

**Démonstration.**
Soit $P$ la matrice de passage $P_{\mathcal{B},\mathcal{B}'}$. Comme les deux bases sont orthonormées, $P \in O(2)$. Supposons d'abord que $\mathcal{B}$ et $\mathcal{B}'$ appartiennent à la même orientation de $E$. Alors $P \in SO(2)$. Or $B = P^{-1}AP$ et $SO(2)$ est un groupe *commutatif*. Donc $B = P^{-1}PA = A$. Supposons maintenant que $\mathcal{B}$ et $\mathcal{B}'$ n'appartiennent pas à la même orientation de $E$. On a donc $\det P = -1$. Introduisons la matrice $J = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Comme $J$ appartient à $O(2)$ et a pour déterminant $-1$, la matrice $Q = PJ$ appartient à $SO(2)$. Or $J^{-1} = J$ donne $P = QJ$. Donc $B = P^{-1}AP = (QJ)^{-1}A(QJ) = JQ^{-1}AQJ = JQ^{-1}QAJ = JAJ$ en utilisant à nouveau le fait que $SO(2)$ est commutatif. Or $A$ est de la forme (7.1). Donc

$$B = JAJ = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} = A^{-1}.$$

□

Il y a donc 2 identifications $SO(E) \cong SO(2)$ possibles. Sélectionner l'une d'entre elles équivaut à sélectionner une orientation de $E$.

Si $E$ est orienté, il y a une identification canonique $SO(E) \cong SO(2)$ : c'est celle qui à $f \in SO(E)$ associe sa matrice $A$ dans une base orthonormée directe quelconque.

### 7.1.3 Notations.

Dans toute la suite de ce chapitre, on suppose $E$ orienté et on fait l'identification canonique $SO(E) \cong SO(2)$ : nous identifions une rotation $f \in SO(E)$ à sa matrice $A$ dans une base orthonormée directe quelconque. La matrice $A$ ne dépend pas de la base orthonormée directe. On notera donc systématiquement $f(\mathbf{v}) = A\mathbf{v}$ pour tout $\mathbf{v} \in E$. Si $\mathcal{B} = (\mathbf{e_1}, \mathbf{e_2})$ est une base orthonormée directe de $E$, on a pour tout vecteur $\mathbf{v} = x\mathbf{e_1} + y\mathbf{e_2} = \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) : f(\mathbf{v}) = A\mathbf{v} = A\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$.

### 7.1.4 Définition de l'angle orienté.

D'après le Lemme 7.1 et l'identification canonique, si $\mathbf{u}, \mathbf{v}$ sont deux vecteurs unitaires de $E$, il existe $A \in SO(2)$ unique tel que $A\mathbf{u} = \mathbf{v}$. □

**Définition 7. 1.** *(i) On appelle angle orienté des deux vecteurs unitaires* $\mathbf{u}, \mathbf{v}$ *du plan euclidien orienté* $E$ *l'unique élément* $A$ *de* $SO(2)$ *tel que* $A\mathbf{u} = \mathbf{v}$.
*(ii) On appelle angle orienté de deux vecteurs non nuls quelconques* $\mathbf{u}, \mathbf{v}$ *de* $E$ *l'angle orienté des deux vecteurs unitaires* $\frac{\mathbf{u}}{\|\mathbf{u}\|}, \frac{\mathbf{v}}{\|\mathbf{v}\|}$.

L'angle orienté de deux vecteurs non nuls $\mathbf{u}, \mathbf{v}$ de $E$ est donc un élément de $SO(2)$. Pour utiliser une notation différente de celle qu'on utilisera plus loin pour ses *déterminations*, on le notera $\widehat{(\mathbf{u}, \mathbf{v})}$.

### 7.1.5 Problème de la mesure des angles orientés.

Mesurer les angles orientés, c'est établir une bijection entre $SO(2)$ et un sous-ensemble de $\mathbb{R}$. Cela va se faire en 2 étapes. D'abord on va montrer que $SO(2)$ est isomorphe au groupe $\mathbb{T}$ des nombres complexes de module 1. Ensuite, on va voir que $\mathbb{T}$ est isomorphe au groupe quotient $\mathbb{R} / 2\pi\mathbb{Z}$ où $\pi$ est un nombre réel que l'on va définir. Cela nous permettra d'associer à un angle des *déterminations* qui sont des nombres réels.

## 7.2 Isomorphisme canonique $\varphi : \mathbb{T} \longrightarrow SO(2)$.

**Lemme 7. 3.** *Soit* $\mathbb{T} = \{z \in \mathbb{C} \,;\, |z| = 1\}$.
*(i) L'application* $\varphi : \mathbb{T} \longrightarrow SO(2)$ *définie par*

$$\varphi(a + ib) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \tag{7.2}$$

*est un isomorphisme du groupe multiplicatif* $\mathbb{T}$ *sur* $SO(2)$.
*(ii) Si l'on identifie l'espace vectoriel* $\mathbb{R}^2$ *au* $\mathbb{R}$*-espace vectoriel* $\mathbb{C}$ *en identifiant le vecteur* $\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) \in \mathbb{R}^2$ *au nombre complexe* $x + iy \in \mathbb{C}$*, l'endomorphisme de* $\mathbb{R}^2$ *dont la matrice dans la base canonique est* $\varphi(a+ib)$ *s'identifie à la multiplication par* $a + ib$ *dans* $\mathbb{C}$ : $w \mapsto (a + ib)w$.

**Démonstration.**

(i) D'après le Th. 2.5, l'application $\varphi$ est bien définie et est une bijection de $\mathbb{T}$ sur $SO(2)$. Il est immédiat de vérifier que c'est un homorphisme de groupes, *i.e.* $\varphi((a + ib)(c + id)) = \varphi(a + ib)\varphi(c + id)$.
(ii) $(a + ib)(x + iy) = ax - by + i(bx + ay) = \left(\begin{smallmatrix} ax-by \\ bx+ay \end{smallmatrix}\right) = \left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right)\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = \varphi(a + ib)\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$.
□

## 7.3 Revêtement universel de $\mathbb{T}$.

### 7.3.1 Fonctions cos et sin.

La fonction exponentielle sur $\mathbb{R}$ $x \mapsto e^x$ est l'application de $\mathbb{R}$ dans $]0,+\infty[$ réciproque de l'application logarithme de $]0,+\infty[$ dans $\mathbb{R}$ définie par $\operatorname{Log} x = \int_1^x \frac{dt}{t}$. On sait que l'on a le développement en série entière

$$e^x = \sum_{n=0}^{+\infty} \frac{x^n}{n!} \quad \forall x \in \mathbb{R}.$$

La série entière $\sum_{n=0}^{+\infty} \frac{x^n}{n!}$ a pour rayon de convergence $R = +\infty$. Cela permet de définir pour tout $z \in \mathbb{C}$ l'exponentielle de $z$ par la formule :

$$e^z = \sum_{n=0}^{+\infty} \frac{z^n}{n!}. \tag{7.3}$$

Pour $z, w \in \mathbb{C}$ fixés, les deux séries $\sum_{n=0}^{+\infty} \frac{z^n}{n!}$ et $\sum_{n=0}^{+\infty} \frac{w^n}{n!}$ sont absolument convergentes, donc la série produit est absolument convergente et sa somme est le produit des sommes :

$$\sum_{n=0}^{+\infty} \left( \sum_{p+q=n} \frac{z^p}{p!} \frac{w^q}{q!} \right) = \left( \sum_{p=0}^{+\infty} \frac{z^p}{p!} \right) \left( \sum_{q=0}^{+\infty} \frac{w^q}{q!} \right) = e^z e^w.$$

Or le terme général $\sum_{p+q=n} \frac{z^p}{p!} \frac{w^q}{q!}$ de la série produit s'écrit

$$\sum_{p=0}^{n} \frac{z^p}{p!} \frac{w^{n-p}}{(n-p)!} = \frac{1}{n!} \sum_{p=0}^{n} C_n^p z^p w^{n-p} = \frac{1}{n!} (z+w)^n$$

d'après la formule du binôme. On a donc

$$e^{z+w} = e^z e^w \quad \forall z, w \in \mathbb{C}. \tag{7.4}$$

En particulier, comme $e^0 = 1$, on a $1 = e^{-z} e^z$, donc $e^z \neq 0$ et $(e^z)^{-1} = e^{-z} \quad \forall z \in \mathbb{C}$. Notons aussi que le conjugué complexe $\overline{e^z}$ est $e^{\bar{z}}$. En effet,

$$\overline{\sum_{n=0}^{+\infty} \frac{z^n}{n!}} = \overline{\lim_{N \to +\infty} \sum_{n=0}^{N} \frac{z^n}{n!}} = \lim_{N \to +\infty} \overline{\sum_{n=0}^{N} \frac{z^n}{n!}} = \lim_{N \to +\infty} \sum_{n=0}^{N} \frac{\bar{z}^n}{n!} = e^{\bar{z}}.$$

En particulier, pour tout $t \in \mathbb{R}$ on a $\overline{e^{it}} = e^{-it}$ *i.e.* $\bar{\zeta} = \zeta^{-1}$ en notant $\zeta = e^{it}$, donc

$$|e^{it}| = 1 \quad \forall t \in \mathbb{R}. \tag{7.5}$$

**Définition 7. 2.** *On définit les fonctions* cos *et* sin *par*

$$\cos t = \mathfrak{Re}\, e^{it}, \quad \sin t = \mathfrak{Im}\, e^{it} \quad \forall t \in \mathbb{R},$$

*où* $\mathfrak{Re}\, z$ *et* $\mathfrak{Im}\, z$ *désignent les parties réelle et imaginaire du nombre complexe* $z$.

(7.5) s'écrit

$$\cos^2 t + \sin^2 t = 1 \quad \forall t \in \mathbb{R}. \tag{7.6}$$

Pour tout $t \in \mathbb{R}$, on a par définition

$$e^{it} = \sum_{n=0}^{+\infty} i^n \frac{t^n}{n!} \tag{7.7}$$

donc

$$\cos t = \sum_{n=0}^{+\infty} (-1)^n \frac{t^{2n}}{(2n)!}, \quad \sin t = \sum_{n=0}^{+\infty} (-1)^n \frac{t^{2n+1}}{(2n+1)!} \tag{7.8}$$

et les séries entières ci-dessus ont pour rayon de convergence $R = +\infty$. En tant que sommes de séries entières, les fonctions $t \mapsto e^{it}$, $t \mapsto \cos t$, $t \mapsto \sin t$ sont de classe $C^\infty$ et les dérivées s'obtiennent par dérivation terme à terme. On a

$$\frac{d}{dt} e^{it} = \sum_{n=1}^{+\infty} i^n n \frac{t^{n-1}}{n!} = i \sum_{k=0}^{+\infty} i^k \frac{t^k}{k!} = ie^{it}. \tag{7.9}$$

En prenant les parties réelles et imaginaires, on obtient

$$\frac{d}{dt} \cos t = -\sin t \,, \quad \frac{d}{dt} \sin t = \cos t \,. \tag{7.10}$$

En particulier, la fonction $t \mapsto e^{it}$ n'est pas constante puisque sa dérivée n'est pas nulle (et même ne s'annule pas) d'après (7.5).

## 7.3.2 Revêtement universel de $\mathbb{T}$.

**Théorème 7. 1.** *Soit $\psi : \mathbb{R} \to \mathbb{T}$ l'application définie par $\psi(t) = e^{it}$.*
*(i) $\psi$ est un homomorphisme de groupes, continu pour les topologies habituelles de $\mathbb{R}$ et $\mathbb{T}$.*
*(ii) Il existe un nombre $\pi > 0$ unique tel que $Ker\, \psi = 2\pi\mathbb{Z}$.*

**Démonstration.**
(i) L'application $\psi$ est définie, puisque $|e^{it}| = 1\ \forall t \in \mathbb{R}$. Ensuite, $\psi$ est un homomorphisme du groupe additif $\mathbb{R}$ dans le groupe multiplicatif $\mathbb{T}$ puisque $\psi(t+s) = e^{i(t+s)} = e^{it}e^{is} = \psi(t)\psi(s)\ \forall t, s \in \mathbb{R}$. De plus $\psi$ est continu puisque $\psi(t) = \cos t + i \sin t\ \forall t \in \mathbb{R}$ par définition des fonctions cos et sin, et que ces fonctions sont continues (par définition de la topologie induite par $\mathbb{R}^2$ sur $\mathbb{T}$, une application de $\mathbb{R}$ dans $\mathbb{T}$ est continue si et seulement si ses deux composantes le sont).
(ii) $\psi$ étant continu, le noyau $H = \operatorname{Ker} \psi = \psi^{-1}(\{1\})$ est un sous-groupe *fermé* de $\mathbb{R}$. On a vu que l'application $t \mapsto e^{it}$ de $\mathbb{R}$ dans $\mathbb{C}$ n'est pas constante, donc $\psi$ n'est pas la constante 1. Cela implique que $H \neq \mathbb{R}$ et par conséquent il existe $a \geqslant 0$ unique tel que $H = a\mathbb{Z}$. Nous allons démontrer que $a > 0$ en montrant que $H \neq \{0\}$.

Pour cela, il suffit de montrer qu'il existe $\beta \neq 0$ tel que $e^{i\beta} = \pm i$. On aura en effet alors $e^{4i\beta} = i^4 = 1$ donc $4\beta \in H$.

Or on a

$$\begin{aligned} e^{i\beta} = \pm i &\Leftrightarrow \cos\beta = 0, \sin\beta = \pm 1 \\ &\Leftrightarrow \cos\beta = 0. \end{aligned}$$

Comme $\cos 0 = 1 > 0$, pour montrer l'existence d'un tel $\beta$, il suffit d'après le théorème des valeurs intermédiaires de montrer que $\cos 2 < 0$, car alors la fonction continue cos s'annulera en au moins un point de l'intervalle ouvert $]0, 2[$. On a ([8], 9.9.5.1, p.215) :

$$\begin{aligned} \cos 2 &= \sum_{n=0}^{\infty} (-1)^n \frac{2^{2n}}{(2n)!} \\ &= 1 + \sum_{k=1}^{\infty} \frac{2^{4k}}{(4k)!} - \sum_{k=0}^{\infty} \frac{2^{4k+2}}{(4k+2)!} \\ &= 1 + \sum_{k=0}^{\infty} \frac{2^{4k+4}}{(4k+4)!} - \sum_{k=0}^{\infty} \frac{2^{4k+2}}{(4k+2)!} \\ &= 1 - \sum_{k=0}^{\infty} \frac{2^{4k+2}}{(4k+2)!} \left(1 - \frac{4}{(4k+4)(4k+3)}\right). \end{aligned}$$

Mais

$$\sum_{k=0}^{\infty} \frac{2^{4k+2}}{(4k+2)!} \left(1 - \frac{4}{(4k+4)(4k+3)}\right) \geqslant \frac{2^2}{2!}\left(1 - \frac{4}{12}\right) = \frac{4}{3} > 1$$

donc $\cos 2 < 0$. On a donc prouvé que $a > 0$.
Maintenant, comme $\operatorname{Ker}\psi = a\mathbb{Z}$ et que $a$ est unique, il existe bien un unique $\pi > 0$ tel que $\operatorname{Ker}\psi = 2\pi\mathbb{Z}$ : il suffit de poser $\pi = \frac{a}{2} > 0$. □

**Corollaire 1.** *(i) Les fonctions* cos *et* sin *sont $2\pi$-périodiques.*
*(ii)* $e^{i\pi} = -1$.
*(iii)* $\sin t = 0 \Leftrightarrow t \in \pi\mathbb{Z}$.
*(iv)* $\cos t = 0 \Leftrightarrow t \in \pi\mathbb{Z} + \frac{\pi}{2}$.
*(v) La fonction* sin *est une bijection strictement croissante de* $[-\frac{\pi}{2}, \frac{\pi}{2}]$ *sur* $[-1, 1]$.
*(vi)* $e^{i\frac{\pi}{2}} = i$.
*(vii) La fonction* cos *est une bijection strictement décroissante de* $[0, \pi]$ *sur* $[-1, 1]$.

**Démonstration.**
(i) résulte immédiatement de $e^{i(t+2\pi)} = e^{2i\pi}e^{it} = e^{it} \, \forall t \in \mathbb{R}$.
(ii) On a $(e^{i\pi})^2 = e^{2i\pi} = 1$ donc $e^{i\pi} = \pm 1$. Or $e^{i\pi} \neq 1$ par définition de $\pi$ puisque $0 < \pi < 2\pi$, donc $e^{i\pi} = -1$.
(iii) $\sin t = 0 \Leftrightarrow e^{it} = \pm 1 \Leftrightarrow e^{2it} = 1 \Leftrightarrow 2t \in 2\pi\mathbb{Z} \Leftrightarrow t \in \pi\mathbb{Z}$.
(iv)

$$\begin{aligned} \cos t = 0 &\Leftrightarrow e^{it} = \pm i \\ &\Leftrightarrow e^{4it} = 1 \text{ et } |\sin t| = 1 \\ &\Leftrightarrow 4t \in 2\pi\mathbb{Z} \text{ et } |\sin t| = 1 \\ &\Leftrightarrow t \in \frac{\pi}{2}\mathbb{Z} \text{ et } |\sin t| = 1 \\ &\Leftrightarrow \exists k \in \mathbb{Z} \quad t = (2k+1)\frac{\pi}{2} \end{aligned}$$

(v) La dérivée de la fonction sin est la fonction cos qui ne s'annule pas sur $]-\frac{\pi}{2},\frac{\pi}{2}[$ et garde donc un signe constant sur cet intervalle. Comme $\cos 0 = 1$, la fonction cos est $> 0$ sur cet intervalle. La fonction sin est donc strictement croissante sur $[-\frac{\pi}{2},\frac{\pi}{2}]$. Comme $\cos\frac{\pi}{2} = 0$, on a $|\sin\frac{\pi}{2}| = 1$, donc $\sin\frac{\pi}{2} = 1$ et $\sin\left(-\frac{\pi}{2}\right) = -1$ puisque la fonction sin est impaire et strictement croissante sur $[-\frac{\pi}{2},\frac{\pi}{2}]$. D'où le résultat.
(vi) On a $\cos\frac{\pi}{2} = 0$ et $\sin\frac{\pi}{2} = 1$.
(vii) La dérivée de la fonction cos est la fonction $-\sin$ qui est $< 0$ sur $]0,\pi[$. La fonction cos est donc strictement décroissante sur $[0,\pi]$. Comme $e^{i\pi} = -1$, on a $\cos\pi = -1$, d'où le résultat. □

**Corollaire 2.** *(i) L'homomorphisme $\psi : \mathbb{R} \to \mathbb{T}$, $\quad \psi(t) = e^{it}$, est surjectif.*
*(ii) Pour tout $z \in \mathbb{T}$, il existe une infinité de $t \in \mathbb{R}$ tels que $z = e^{it}$.*
*(iii) Pour $t, s \in \mathbb{R}$, on a $e^{it} = e^{is} \Leftrightarrow t - s \in 2\pi\mathbb{Z}$.*
*(iv) Pour tout $z \in \mathbb{T}$, il existe $s \in ]-\pi,\pi]$ unique tel que $z = e^{is}$.*
*(v) L'application $\psi$ définit par passage au quotient un isomorphisme*

$$\widetilde{\psi} : \mathbb{R} \,/\, 2\pi\mathbb{Z} \to \mathbb{T} \tag{7.11}$$

*du groupe additif $\mathbb{R} \,/\, 2\pi\mathbb{Z}$ sur le groupe multiplicatif $\mathbb{T}$ tel que*

$$\widetilde{\psi}([t]) = \psi(t) = e^{it} \quad \forall t \in \mathbb{R}.$$

□

**Démonstration.**

(i) Soit $z = x + iy \in \mathbb{T}$. On a $|z|^2 = x^2 + y^2 = 1$, donc $y \in [-1,1]$ et, d'après le Corollaire 1, il existe $t \in [-\frac{\pi}{2},\frac{\pi}{2}]$ tel que $y = \sin t$. Alors $x^2 = 1 - y^2 = 1 - \sin^2 t = \cos^2 t$ donc $x = \pm\cos t$. Si $x = \cos t$, on a directement $z = e^{it}$. Si $x = -\cos t$, on a

$$z = -\cos t + i\sin t = -(\cos t - i\sin t) = -e^{-it} = e^{i\pi}e^{-it} = e^{i(\pi - t)}.$$

(ii) et (iii) : Soit $z \in \mathbb{T}$. D'après (i), il existe $t$ tel que $z = e^{it}$. Alors

$$z = e^{is} \Leftrightarrow e^{it} = e^{is} \Leftrightarrow e^{i(t-s)} = 1 \Leftrightarrow t - s \in \operatorname{Ker}\psi.$$

D'où le résultat puisque $\operatorname{Ker}\psi = 2\pi\mathbb{Z}$.
(iv) Soit $z \in \mathbb{T}$. D'après (ii) et (iii), il existe $t \in \mathbb{R}$ tel que $z = e^{it}$ et l'ensemble $E_z$ des $s \in \mathbb{R}$ tels que $z = e^{is}$ est l'ensemble des réels de la forme $s = t + 2k\pi$, $k \in \mathbb{Z}$. Or si $s_1 = t + 2k_1\pi$ et $s_2 = t + 2k_2\pi$ $(k_1, k_2 \in \mathbb{Z})$ sont deux éléments distincts de $E_z$, on a $k_1 \neq k_2$ et $|s_1 - s_2| = 2|k_1 - k_2|\pi \geqslant 2\pi$. Donc il existe *au plus un* élément $s \in E_z$ tel que $s \in ]-\pi,\pi]$. Par ailleurs, pour tout $N \in \mathbb{N}^*$ on a

$$](-2N-1)\pi, (2N+1)\pi] = \bigcup_{k=-N}^{k=N} ](2k-1)\pi, (2k+1)\pi]$$

donc

$$\mathbb{R} = \bigcup_{k\in\mathbb{Z}} ](2k-1)\pi, (2k+1)\pi].$$

Il existe donc un $k \in \mathbb{Z}$ tel que $t \in ](2k-1)\pi, (2k+1)\pi]$ et alors $s = t - 2k\pi \in ]-\pi,\pi]$. Ainsi $E_z \cap ]-\pi,\pi] = \{s\}$.

(v) $\psi$ est un homorphisme surjectif de $\mathbb{R}$ sur $\mathbb{T}$, et son noyau est $2\pi\mathbb{Z}$. Par décomposition canonique de $\psi$, on obtient donc un isomorphisme $\widetilde{\psi} : \mathbb{R} \,/\, 2\pi\mathbb{Z} \to \mathbb{T}$. □

**Corollaire 3.** *(i) L'application $R = \varphi \circ \psi : \mathbb{R} \to SO(2)$ (où $\varphi : \mathbb{T} \longrightarrow SO(2)$ est l'application définie par par 7.2) est un homorphisme surjectif du groupe additif $\mathbb{R}$ sur le groupe $SO(2)$, de noyau $Ker R = 2\pi\mathbb{Z}$. On a pour $t \in \mathbb{R}$ :*

$$R(t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}. \tag{7.12}$$

*(ii) Pour tout $A \in SO(2)$, il existe une infinité de $t \in \mathbb{R}$ tels que $A = R(t)$.*
*(iii) Pour $t, s \in \mathbb{R}$, on a $R(t) = R(s) \Leftrightarrow t - s \in 2\pi\mathbb{Z}$.*
*(iv) Pour tout $A \in SO(2)$, il existe $t \in ]-\pi, \pi]$ unique tel que $A = R(t)$.*
*(v) L'application $R$ définit par passage au quotient un isomorphisme*

$$\widetilde{R} : \mathbb{R} \,/\, 2\pi\mathbb{Z} \to SO(2). \tag{7.13}$$

*du groupe additif $\mathbb{R} \,/\, 2\pi\mathbb{Z}$ sur le groupe multiplicatif $SO(2)$ tel que*

$$\widetilde{R}([t]) = R(t) \quad \forall t \in \mathbb{R}.$$

**Démonstration.**

(i) $\psi$ est un homorphisme surjectif de $\mathbb{R}$ sur $\mathbb{T}$ et $\varphi$ est un isomorphisme de $\mathbb{T}$ sur $SO(2)$. Donc $R = \varphi \circ \psi$ est un homorphisme surjectif de $\mathbb{R}$ sur $SO(2)$. Par définition de $\varphi$, on a pour tout $t \in \mathbb{R}$

$$R(t) = \varphi(\psi(t)) = \varphi(e^{it}) = \varphi(\cos t + i \sin t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}.$$

Enfin

$$\begin{aligned} R(t) = I \quad &\Leftrightarrow \quad \varphi(\psi(t)) = I \\ &\Leftrightarrow \quad \psi(t) = 1 \quad \text{(car } \varphi \text{ est un isomorphisme)} \\ &\Leftrightarrow \quad t \in 2\pi\mathbb{Z}. \end{aligned}$$

(ii), (iii) et (iv). Pour tout $t \in \mathbb{R}$ on a $A = R(t) \Leftrightarrow A = \varphi(\psi(t)) \Leftrightarrow z = \psi(t)$ en posant $z = \varphi^{-1}(A)$. D'où le résultat d'après le Coroll.2 (ii), (iii), (iv).
(v) Résulte immédiatement de (i) par décomposition de $R$. □

**Définition 7. 3.** *L'application $\psi : \mathbb{R} \to \mathbb{T}$ définie par $\psi(t) = e^{it}$ et introduite au Théorème 7.1 est appelée le revêtement universel de $\mathbb{T}$.*

## 7.4 Déterminations de l'angle orienté de 2 vecteurs.

**Définition 7. 4.** *(i) Soient $\mathbf{u}, \mathbf{v}$ deux vecteurs unitaires du plan euclidien orienté $E$, et $A = \widehat{(\mathbf{u}, \mathbf{v})}$ l'angle orienté des deux vecteurs $\mathbf{u}, \mathbf{v}$, i.e. l'unique élément $A$ de $SO(2)$ tel que $A\mathbf{u} = \mathbf{v}$. On appelle détermination de la mesure en radians de l'angle*

*orienté, ou plus simplement détermination de l'angle orienté* $\widehat{\widehat{(\mathbf{u},\mathbf{v})}}$, *tout réel* $t \in \mathbb{R}$ *tel que* $A = R(t)$ *où* $R(t)$ *est défini par (7.12).*
*(ii) On appelle détermination de l'angle orienté de deux vecteurs non nuls quelconques* $\mathbf{u}, \mathbf{v}$ *de* $E$ *une détermination de l'angle orienté des deux vecteurs unitaires* $\frac{\mathbf{u}}{\|\mathbf{u}\|}, \frac{\mathbf{v}}{\|\mathbf{v}\|}$.

Par abus de langage, pour $t \in \mathbb{R}$, l'endomorphisme de l'espace vectoriel orienté $E$ dont la matrice dans la base directe $\mathcal{B}$ est $R(t)$ sera appelée rotation d' "angle" $t$.

D'après le Corollaire 3 du Théorème 7.1, l'angle orienté $\widehat{\widehat{(\mathbf{u},\mathbf{v})}}$ possède une infinité de déterminations, et deux déterminations diffèrent d'un multiple de $2\pi$.

**Définition 7. 5.** *On appelle détermination principale de l'angle orienté* $\widehat{\widehat{(\mathbf{u},\mathbf{v})}}$ *l'unique détermination* $t_0$ *telle que* $t_0 \in ]-\pi, \pi]$.

Si $t \in \mathbb{R}$ est une détermination de $\widehat{\widehat{(\mathbf{u},\mathbf{v})}}$, on notera $t = \widehat{(\mathbf{u},\mathbf{v})}$.

Si $t = \widehat{(\mathbf{u},\mathbf{v})}$ est une détermination de l'angle $\widehat{\widehat{(\mathbf{u},\mathbf{v})}}$, la classe d'équivalence de $t$ modulo $2\pi\mathbb{Z}$ est l'unique élément $\mathcal{C}_{\mathbf{u},\mathbf{v}}$ de $\mathbb{R}\,/\,2\pi\mathbb{Z}$ tel que $\widetilde{R}(\mathcal{C}_{\mathbf{u},\mathbf{v}}) = \widehat{\widehat{(\mathbf{u},\mathbf{v})}}$.

On identifiera cette classe à l'angle $\widehat{\widehat{(\mathbf{u},\mathbf{v})}}$ par l'isomorphisme $\widetilde{R}$.

Cela pose un problème de notation, car le groupe $\mathbb{R}\,/\,2\pi\mathbb{Z}$ est un groupe *additif* alors que le groupe $SO(2)$ est un groupe *multiplicatif.* L'habitude fait que l'on opte pour la notation additive, *i.e.* l'addition dans $\mathbb{R}\,/\,2\pi\mathbb{Z}$ .

**Proposition 7. 1.** *(i)*

$$\widehat{\widehat{(-\mathbf{u},-\mathbf{v})}} = \widehat{\widehat{(\mathbf{u},\mathbf{v})}} \tag{7.14}$$

*ou, en prenant des déterminations :*

$$\widehat{(-\mathbf{u},-\mathbf{v})} \equiv \widehat{(\mathbf{u},\mathbf{v})} \pmod{2\pi}. \tag{7.15}$$

*(ii)*

$$\widehat{\widehat{(\mathbf{v},\mathbf{u})}} = -\widehat{\widehat{(\mathbf{u},\mathbf{v})}} \tag{7.16}$$

*ou, en prenant des déterminations :*

$$\widehat{(\mathbf{v},\mathbf{u})} \equiv -\widehat{(\mathbf{u},\mathbf{v})} \pmod{2\pi}. \tag{7.17}$$

**Démonstration.**

On peut supposer $\mathbf{u}, \mathbf{v}$ unitaires. Soit $t$ une détermination de l'angle orienté $\widehat{\widehat{(\mathbf{u},\mathbf{v})}}$.
(i) On a $\mathbf{v} = R(t)\mathbf{u}$, donc $-\mathbf{v} = R(t)(-\mathbf{u})$. Ainsi $t$ est aussi une détermination de $\widehat{\widehat{(-\mathbf{u},-\mathbf{v})}}$, ce qui équivaut à dire que $\widehat{\widehat{(-\mathbf{u},-\mathbf{v})}} = \widehat{\widehat{(\mathbf{u},\mathbf{v})}}$.
(ii) On a $\mathbf{u} = (R(t))^{-1}\mathbf{v} = R(-t)\mathbf{v}$, donc $-t$ une détermination de $\widehat{\widehat{(\mathbf{v},\mathbf{u})}}$, d'où $\widehat{\widehat{(\mathbf{v},\mathbf{u})}} = -\widehat{\widehat{(\mathbf{u},\mathbf{v})}}$. □

## 7.5 Additivité des angles.

**Théorème 7. 2.** *Soient* $\mathbf{u}, \mathbf{v}, \mathbf{w}$ *trois vecteurs non nuls. Alors*

$$\widehat{\widehat{(\mathbf{u}, \mathbf{w})}} = \widehat{\widehat{(\mathbf{u}, \mathbf{v})}} + \widehat{\widehat{(\mathbf{v}, \mathbf{w})}} \tag{7.18}$$

*ou, en prenant des déterminations :*

$$\widehat{(\mathbf{u}, \mathbf{w})} \equiv \widehat{(\mathbf{u}, \mathbf{v})} + \widehat{(\mathbf{v}, \mathbf{w})} \qquad (\text{mod } 2\pi). \tag{7.19}$$

**Démonstration.**

On peut supposer $\mathbf{u}, \mathbf{v}$ unitaires. Soit $t$ une détermination de $\widehat{\widehat{(\mathbf{u}, \mathbf{v})}}$ et $s$ une détermination de $\widehat{\widehat{(\mathbf{v}, \mathbf{w})}}$. Alors $\mathbf{v} = R(t)\mathbf{u}$ et $\mathbf{w} = R(s)\mathbf{v}$, donc

$$\mathbf{w} = R(s)R(t)\mathbf{u} = R(s+t)\mathbf{u}$$

*i.e.* $s + t$ est une détermination de $\widehat{\widehat{(\mathbf{u}, \mathbf{w})}}$. □

**Corollaire.** *Soient* $\mathbf{u}, \mathbf{v}, \mathbf{w}$ *des vecteurs non nuls,* $t = \widehat{(\mathbf{u}, \mathbf{v})}$ *et* $s = \widehat{(\mathbf{u}, \mathbf{w})}$ *des déterminations de* $\widehat{\widehat{(\mathbf{u}, \mathbf{v})}}$ *et* $\widehat{\widehat{(\mathbf{u}, \mathbf{w})}}$*. Alors* $s - t$ *est une détermination de* $\widehat{\widehat{(\mathbf{v}, \mathbf{w})}}$*.*

**Démonstration.**

Soit $\widehat{(\mathbf{v}, \mathbf{w})}$ une détermination de l'angle $\widehat{\widehat{(\mathbf{v}, \mathbf{w})}}$. D'après le Th. 7.2, on a $s \equiv t + \widehat{(\mathbf{v}, \mathbf{w})} \pmod{2\pi}$. Donc $\widehat{(\mathbf{v}, \mathbf{w})} \equiv s - t \pmod{2\pi}$. Alors $s - t$ est aussi une détermination de $\widehat{\widehat{(\mathbf{v}, \mathbf{w})}}$. □

## 7.6 Formules.

**Théorème 7. 3.** *Soient* $\mathbf{u}, \mathbf{v}$ *deux vecteurs non nuls, et* $t = \widehat{(\mathbf{u}, \mathbf{v})}$ *une détermination de l'angle* $\widehat{\widehat{(\mathbf{u}, \mathbf{v})}}$*. Alors :*

$$(\mathbf{u}|\mathbf{v}) = \|\mathbf{u}\|\|\mathbf{v}\| \cos t \tag{7.20}$$
$$\det_{\mathcal{B}}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u}\|\|\mathbf{v}\| \sin t \tag{7.21}$$

*où* $\mathcal{B}$ *est une base orthonormée directe quelquonque de* $E$*.*

**Démonstration.**

Par définition de $t$, $\frac{\mathbf{v}}{\|\mathbf{v}\|} = R(t)\frac{\mathbf{u}}{\|\mathbf{u}\|}$, donc

$$\mathbf{v} = \frac{\| \mathbf{v} \|}{\| \mathbf{u} \|} R(t)\mathbf{u}.$$

Dans la base $\mathcal{B}$, on a :

$$\mathbf{u} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \ \|\mathbf{u}\| = \sqrt{\alpha^2 + \beta^2},$$

donc

$$\mathbf{v} = \frac{\| \mathbf{v} \|}{\| \mathbf{u} \|} \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \frac{\| \mathbf{v} \|}{\| \mathbf{u} \|} \begin{pmatrix} \alpha \cos t - \beta \sin t \\ \alpha \sin t + \beta \cos t \end{pmatrix},$$

$$(\mathbf{u}|\mathbf{v}) = \frac{\|\mathbf{v}\|}{\|\mathbf{u}\|}(\alpha^2+\beta^2)\cos t = \|\mathbf{u}\|\|\mathbf{v}\|\cos t\,.$$

Enfin,

$$\begin{aligned}\det_{\mathcal{B}}(\mathbf{u},\mathbf{v}) &= \frac{\|\mathbf{v}\|}{\|\mathbf{u}\|}\begin{vmatrix}\alpha & \alpha\cos t-\beta\sin t\\ \beta & \alpha\sin t+\beta\cos t\end{vmatrix}\\ &= \frac{\|\mathbf{v}\|}{\|\mathbf{u}\|}(\alpha^2+\beta^2)\sin t\\ &= \|\mathbf{u}\|\|\mathbf{v}\|\sin t\,.\end{aligned}$$

□

## 7.7 Somme des angles d'un triangle.

**Théorème 7. 4.** *Soient $A, B, C$ trois points non alignés du plan affine euclidien orienté, et (voir Fig. 7.1)*

$$(\widehat{\mathbf{AB},\mathbf{AC}}),\ (\widehat{\mathbf{BC},\mathbf{BA}}),\ (\widehat{\mathbf{CA},\mathbf{CB}})$$

*les déterminations principales des angles orientés*

$$(\widehat{\widehat{\mathbf{AB},\mathbf{AC}}}),\ (\widehat{\widehat{\mathbf{BC},\mathbf{BA}}}),\ (\widehat{\widehat{\mathbf{CA},\mathbf{CB}}})$$

*respectivement. Alors*
*(i)*

$$(\widehat{\mathbf{AB},\mathbf{AC}}) + (\widehat{\mathbf{BC},\mathbf{BA}}) + (\widehat{\mathbf{CA},\mathbf{CB}}) = \pm\pi. \qquad (7.22)$$

*(ii)* $(\widehat{\mathbf{AB},\mathbf{AC}}), (\widehat{\mathbf{BC},\mathbf{BA}}), (\widehat{\mathbf{CA},\mathbf{CB}})$ *sont tous les trois du même signe.*
*(iii)*

$$\|\mathbf{AC}\| = \|\mathbf{BC}\| \Leftrightarrow (\widehat{\mathbf{AB},\mathbf{AC}}) = (\widehat{\mathbf{BC},\mathbf{BA}})\,. \qquad (7.23)$$

**Démonstration.**

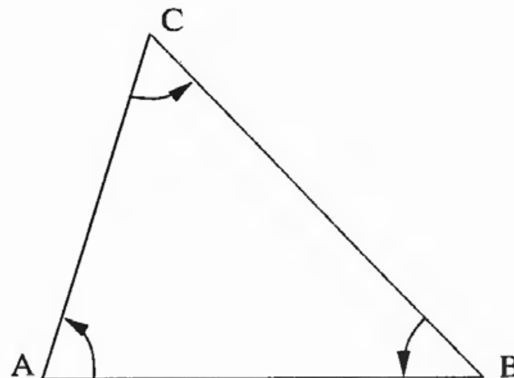



FIG. 7.1: *Somme des angles d'un triangle.*

(i) D'après la propriété d'additivité des angles, on a en prenant les déterminations principales

$$(\widehat{\mathbf{AB},\mathbf{AC}}) + (\widehat{\mathbf{AC},\mathbf{BC}}) + (\widehat{\mathbf{BC},\mathbf{BA}}) \equiv (\widehat{\mathbf{AB},\mathbf{BA}}) \pmod{2\pi}.$$

D'après la Proposition 7.1 , toujours avec la détermination principale,

$$(\widehat{\mathbf{AC},\mathbf{BC}}) = (\widehat{-\mathbf{AC},-\mathbf{BC}}) = (\widehat{\mathbf{CA},\mathbf{CB}}).$$

De plus $(\widehat{\mathbf{AB},\mathbf{BA}}) = (\widehat{\mathbf{AB},-\mathbf{AB}}) = \pi$ donc

$$(\widehat{\mathbf{AB},\mathbf{AC}}) + (\widehat{\mathbf{BC},\mathbf{BA}}) + (\widehat{\mathbf{CA},\mathbf{CB}}) \equiv \pi \pmod{2\pi}.$$

*i.e.*

$$(\widehat{\mathbf{AB},\mathbf{AC}}) + (\widehat{\mathbf{BC},\mathbf{BA}}) + (\widehat{\mathbf{CA},\mathbf{CB}}) = \pi + 2k\pi, \quad k \in \mathbb{Z}. \tag{7.24}$$

Mais les trois points $A, B, C$ n'étant pas alignés, on a

$$|(\widehat{\mathbf{AB},\mathbf{AC}})| < \pi,\ |(\widehat{\mathbf{BC},\mathbf{BA}})| < \pi,\ |(\widehat{\mathbf{CA},\mathbf{CB}})| < \pi$$

donc

$$|(\widehat{\mathbf{AB},\mathbf{AC}}) + (\widehat{\mathbf{BC},\mathbf{BA}}) + (\widehat{\mathbf{CA},\mathbf{CB}})| < 3\pi\,.$$

Par conséquent dans l'équation (7.24) on a nécessairement $k = 0$ ou $k = -1$, ce qui donne (7.22).

(ii) On a par définition d'une détermination

$$\frac{\mathbf{AC}}{\|\mathbf{AC}\|} = R((\widehat{\mathbf{AB},\mathbf{AC}}))\frac{\mathbf{AB}}{\|\mathbf{AB}\|}.$$

Utilisons la base orthonormée directe $(\mathbf{u},\mathbf{v})$, avec $\mathbf{u} = \frac{\mathbf{AB}}{\|\mathbf{AB}\|}$ et $\mathbf{v}$ vecteur directement perpendiculaire. Dans cette base,

$$\begin{aligned}
\mathbf{AC} &= \|\mathbf{AC}\| \begin{pmatrix} \cos(\widehat{\mathbf{AB},\mathbf{AC}}) & -\sin(\widehat{\mathbf{AB},\mathbf{AC}}) \\ \sin(\widehat{\mathbf{AB},\mathbf{AC}}) & \cos(\widehat{\mathbf{AB},\mathbf{AC}}) \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
&= \|\mathbf{AC}\| \begin{pmatrix} \cos(\widehat{\mathbf{AB},\mathbf{AC}}) \\ \sin(\widehat{\mathbf{AB},\mathbf{AC}}) \end{pmatrix}
\end{aligned}$$

De même,

$$\frac{\mathbf{BC}}{\|\mathbf{BC}\|} = R((\widehat{\mathbf{BA},\mathbf{BC}}))\frac{\mathbf{BA}}{\|\mathbf{BA}\|} = R((\widehat{\mathbf{BA},\mathbf{BC}}))(-\mathbf{u})$$

donne

$$\begin{aligned}
\mathbf{BC} &= \|\mathbf{BC}\| \begin{pmatrix} \cos(\widehat{\mathbf{BA},\mathbf{BC}}) & -\sin(\widehat{\mathbf{BA},\mathbf{BC}}) \\ \sin(\widehat{\mathbf{BA},\mathbf{BC}}) & \cos(\widehat{\mathbf{BA},\mathbf{BC}}) \end{pmatrix} \begin{pmatrix} -1 \\ 0 \end{pmatrix} \\
&= \|\mathbf{BC}\| \begin{pmatrix} -\cos(\widehat{\mathbf{BA},\mathbf{BC}}) \\ -\sin(\widehat{\mathbf{BA},\mathbf{BC}}) \end{pmatrix} \\
&= \|\mathbf{BC}\| \begin{pmatrix} -\cos(\widehat{\mathbf{BC},\mathbf{BA}}) \\ \sin(\widehat{\mathbf{BC},\mathbf{BA}}) \end{pmatrix}
\end{aligned}$$

puisque $(\widehat{\mathbf{BA},\mathbf{BC}}) = -(\widehat{\mathbf{BC},\mathbf{BA}})$. Donc

$$\begin{aligned}\mathbf{AB} = \mathbf{AC} + \mathbf{CB} &= \|\mathbf{AC}\| \begin{pmatrix} \cos(\widehat{\mathbf{AB},\mathbf{AC}}) \\ \sin(\widehat{\mathbf{AB},\mathbf{AC}}) \end{pmatrix} - \|\mathbf{BC}\| \begin{pmatrix} -\cos(\widehat{\mathbf{BC},\mathbf{BA}}) \\ \sin(\widehat{\mathbf{BC},\mathbf{BA}}) \end{pmatrix} \\ &= \begin{pmatrix} \|\mathbf{AC}\|\cos(\widehat{\mathbf{AB},\mathbf{AC}}) + \|\mathbf{BC}\|\cos(\widehat{\mathbf{BC},\mathbf{BA}}) \\ \|\mathbf{AC}\|\sin(\widehat{\mathbf{AB},\mathbf{AC}}) - \|\mathbf{BC}\|\sin(\widehat{\mathbf{BC},\mathbf{BA}}) \end{pmatrix}\end{aligned}$$

*i.e.*

$$\begin{pmatrix} \|\mathbf{AB}\| \\ 0 \end{pmatrix} = \begin{pmatrix} \|\mathbf{AC}\|\cos(\widehat{\mathbf{AB},\mathbf{AC}}) + \|\mathbf{BC}\|\cos(\widehat{\mathbf{BC},\mathbf{BA}}) \\ \|\mathbf{AC}\|\sin(\widehat{\mathbf{AB},\mathbf{AC}}) - \|\mathbf{BC}\|\sin(\widehat{\mathbf{BC},\mathbf{BA}}) \end{pmatrix}. \tag{7.25}$$

Cela implique que

$$\|\mathbf{AC}\|\sin(\widehat{\mathbf{AB},\mathbf{AC}}) - \|\mathbf{BC}\|\sin(\widehat{\mathbf{BC},\mathbf{BA}}) = 0,$$

donc $\|\mathbf{AC}\|\sin(\widehat{\mathbf{AB},\mathbf{AC}}) = \|\mathbf{BC}\|\sin(\widehat{\mathbf{BC},\mathbf{BA}})$ et ainsi $(\widehat{\mathbf{AB},\mathbf{AC}})$ et $(\widehat{\mathbf{BC},\mathbf{BA}})$ sont de même signe. Par permutation circulaire de $A, B, C$, le résultat en découle.
(iii) D'après l'équation (7.25), la condition $\|\mathbf{AC}\| = \|\mathbf{BC}\|$ implique

$$\sin(\widehat{\mathbf{AB},\mathbf{AC}}) = \sin(\widehat{\mathbf{BC},\mathbf{BA}}) \tag{7.26}$$

et

$$\|\mathbf{AB}\| = \|\mathbf{AC}\|(\cos(\widehat{\mathbf{AB},\mathbf{AC}}) + \cos(\widehat{\mathbf{BC},\mathbf{BA}})),$$

donc en particulier

$$\cos(\widehat{\mathbf{AB},\mathbf{AC}}) \neq -\cos(\widehat{\mathbf{BC},\mathbf{BA}}). \tag{7.27}$$

La conjonction des deux équations (7.26) et (7.27) implique $(\widehat{\mathbf{AB},\mathbf{AC}}) = (\widehat{\mathbf{BC},\mathbf{BA}})$.

Réciproquement, si cette condition est satisfaite, l'équation (7.25) montre que $\|\mathbf{AC}\| = \|\mathbf{BC}\|$. □

**Définition 7. 6.** *Un triangle est dit isocèle s'il possède deux côtés de même longueur.*

## 7.8 Angle au centre et angle inscrit.

**Théorème 7. 5.** *Soient $A, B, C$ trois points non alignés du plan affine euclidien orienté, et $I$ le centre du cercle passant par ces trois points. Alors*

$$(\widehat{\widehat{\mathbf{IB},\mathbf{IC}}}) = 2\,(\widehat{\widehat{\mathbf{AB},\mathbf{AC}}})$$

*ou encore en prenant des déterminations :*

$$(\widehat{\mathbf{IB},\mathbf{IC}}) \equiv 2\,(\widehat{\mathbf{AB},\mathbf{AC}}) \qquad (\mathrm{mod}\ 2\pi).$$

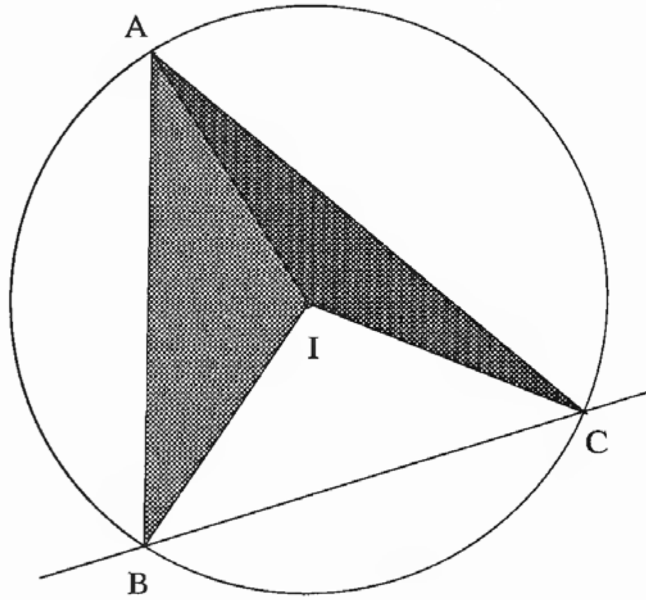



FIG. 7.2: *Angle au centre et angle inscrit (cas 1).*

**Démonstration.**

Le triangle $IAB$ étant isocèle, (voir Fig. 7.2 ou Fig. 7.3, la démonstration étant valable pour les deux cas) on a :

$$(\widehat{\mathbf{AB},\mathbf{AI}}) \equiv (\widehat{\mathbf{BI},\mathbf{BA}}) \pmod{2\pi}.$$

Or d'après le Théorème 7.4,

$$(\widehat{\mathbf{AB},\mathbf{AI}}) + (\widehat{\mathbf{IA},\mathbf{IB}}) + (\widehat{\mathbf{BI},\mathbf{BA}}) \equiv \pi \pmod{2\pi}.$$

Donc

$$2(\widehat{\mathbf{AB},\mathbf{AI}}) + (\widehat{\mathbf{IA},\mathbf{IB}}) \equiv \pi \pmod{2\pi}. \tag{7.28}$$

Le triangle $IAC$ étant aussi isocèle, on obtient de même :

$$2(\widehat{\mathbf{AC},\mathbf{AI}}) + (\widehat{\mathbf{IA},\mathbf{IC}}) \equiv \pi \pmod{2\pi}. \tag{7.29}$$

Mais, d'après la propriété d'additivité des angles, on a

$$(\widehat{\mathbf{AB},\mathbf{AC}}) \equiv (\widehat{\mathbf{AB},\mathbf{AI}}) + (\widehat{\mathbf{AI},\mathbf{AC}}) \pmod{2\pi},$$

ce qui donne en utilisant (7.28) et (7.29) :

$$\begin{aligned}
2(\widehat{\mathbf{AB},\mathbf{AC}}) &\equiv \pi - (\widehat{\mathbf{IA},\mathbf{IB}}) - (\pi - (\widehat{\mathbf{IA},\mathbf{IC}})) \pmod{2\pi} \\
&\equiv (\widehat{\mathbf{IA},\mathbf{IC}}) - (\widehat{\mathbf{IA},\mathbf{IB}}) \pmod{2\pi} \\
&\equiv (\widehat{\mathbf{IB},\mathbf{IC}}) \pmod{2\pi}.
\end{aligned}$$

□

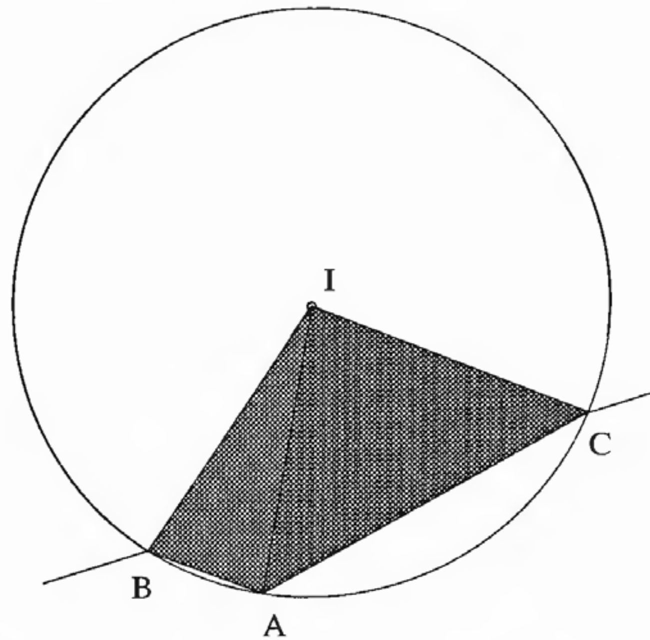



FIG. 7.3: *Angle au centre et angle inscrit (cas 2).*

## 7.9 Arc capable, cocyclicité.

**Théorème 7. 6.** *Soient $B, C$ deux points distincts du plan affine euclidien orienté, et $\alpha \in ]0, \pi[$. L'ensemble des points $M \neq B, C$ tels que*

$$(\widehat{\mathbf{MB}, \mathbf{MC}}) \equiv \alpha \qquad (\text{mod } \pi) \tag{7.30}$$

*est un cercle passant par $B, C$, privé des points $B, C$. La corde $BC$ sépare ce cercle en deux arcs, dont l'un est l'ensemble des points $M \neq B, C$ tels que $(\widehat{\mathbf{MB}, \mathbf{MC}}) \equiv \alpha$ $(\text{mod } 2\pi)$, et l'autre l'ensemble des points $M \neq B, C$ tels que $(\widehat{\mathbf{MB}, \mathbf{MC}}) \equiv \alpha - \pi$ $(\text{mod } 2\pi)$ (voir Fig. 7.4).*

**Démonstration.**

Soit $\Delta$ la médiatrice du segment $[B, C]$, et $I$ un point quelconque de $\Delta$. Si $I$ n'est pas le milieu de $[B, C]$, on a dans le triangle $IBC$ d'après (7.22) :

$$(\widehat{\mathbf{BC}, \mathbf{BI}}) + (\widehat{\mathbf{IB}, \mathbf{IC}}) + (\widehat{\mathbf{CI}, \mathbf{CB}}) \equiv \pi \qquad (\text{mod } 2\pi),$$

en notant que dans (7.22) il s'agit des déterminations principales, alors qu'ici il s'agit de déterminations quelconques. Comme le triangle est isocèle, cela s'écrit :

$$2(\widehat{\mathbf{BC}, \mathbf{BI}}) + (\widehat{\mathbf{IB}, \mathbf{IC}}) \equiv \pi \qquad (\text{mod } 2\pi).$$

Cette relation est encore vraie si $I$ est le milieu de $[B, C]$. Elle est donc vraie pour tout $I \in \Delta$. On a alors pour tout $I \in \Delta$ :

$$(\widehat{\mathbf{IB}, \mathbf{IC}}) \equiv 2\alpha \quad (\text{mod } 2\pi) \;\Leftrightarrow\; 2(\widehat{\mathbf{BC}, \mathbf{BI}}) \equiv \pi - 2\alpha \quad (\text{mod } 2\pi) \tag{7.31}$$

$$\Leftrightarrow\; (\widehat{\mathbf{BC}, \mathbf{BI}}) \equiv \frac{\pi}{2} - \alpha \quad (\text{mod } \pi). \tag{7.32}$$

Considérons maintenant la droite $\mathcal{D}_\alpha$ passant par $B$ et ayant pour vecteur directeur le vecteur unitaire $\mathbf{u}_\alpha$ défini par $(\widehat{\mathbf{BC}, \mathbf{u}_\alpha}) \equiv \frac{\pi}{2} - \alpha \quad (\text{mod } 2\pi)$.

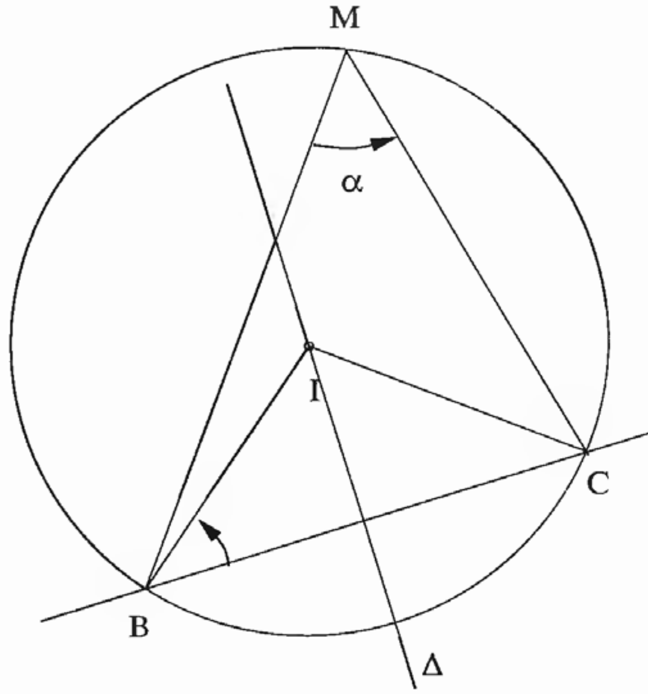



FIG. 7.4: *Arc capable.*

Un point $M \neq B$ du plan appartient à $\mathcal{D}_\alpha$ si et seulement si

$$(\widehat{\mathbf{BC},\mathbf{BM}}) \equiv \frac{\pi}{2} - \alpha \pmod{2\pi} \text{ ou } (\widehat{\mathbf{BC},\mathbf{BM}}) \equiv \pi + \frac{\pi}{2} - \alpha \pmod{2\pi},$$

*i.e.*

$$\mathcal{D}_\alpha = \{B\} \cup \{M\,;\, M \neq B\,;\, (\widehat{\mathbf{BC},\mathbf{BM}}) \equiv \frac{\pi}{2} - \alpha \pmod{\pi}\}. \tag{7.33}$$

En comparant (7.32) et (7.33) , on voit qu'il existe donc un point unique $I_\alpha \in \Delta$ tel que $(\widehat{\mathbf{IB},\mathbf{IC}}) \equiv 2\alpha \pmod{2\pi}$, et que $I_\alpha$ est le point d'intersection de $\Delta$ avec $\mathcal{D}_\alpha$.

Or si $M$ est un point quelconque du plan, non aligné avec $B$ et $C$, on a d'après le Théorème 7.5, $(\widehat{\mathbf{MB},\mathbf{MC}}) \equiv \alpha \pmod{\pi}$ si et seulement si le centre $I$ du cercle passant par les 3 points $M, B, C$ vérifie

$$(\widehat{\mathbf{IB},\mathbf{IC}}) \equiv 2\alpha \pmod{2\pi},$$

*i.e.* $I = I_\alpha$.

L'ensemble des points $M \neq B, C$ vérifiant (7.30) est donc le cercle de centre $I_\alpha$ et passant par les points $B$ et $C$, privé des points $B$ et $C$.

De plus, la corde $BC$ sépare ce cercle en deux arcs. Soit $t$ la détermination principale de l'angle $(\widehat{\widehat{\mathbf{MB},\mathbf{MC}}})$.

Sur l'un des deux arcs, on a $t > 0$, donc $t \in ]0, \pi[$. Or $t \equiv \alpha \pmod{\pi}$ d'après (7.30). Donc $t = \alpha$.

Sur l'autre arc, on a $t < 0$, donc $t \in ]-\pi, 0[$. Or $\alpha - \pi \in ]-\pi, 0[$ et $t \equiv \alpha - \pi \pmod{\pi}$ d'après (7.30). Donc $t = \alpha - \pi$. □

**Corollaire.** *Soient $A, B, C, D$ quatre points deux-à-deux distincts du plan affine euclidien orienté. Pour que les quatre points $A, B, C, D$ soient alignés ou cocycliques, il faut et il suffit que*

$$(\widehat{\mathbf{AB},\mathbf{AC}}) \equiv (\widehat{\mathbf{DB},\mathbf{DC}}) \pmod{\pi} \tag{7.34}$$

**Démonstration.**

D'après le Théorème 7.6, les quatre points $A, B, C, D$ sont cocycliques si et seulement si il existe $\alpha \in ]0, \pi[$ tel que $(\widehat{\mathbf{AB}, \mathbf{AC}}) \equiv \alpha \pmod{\pi}$ et $(\widehat{\mathbf{DB}, \mathbf{DC}}) \equiv \alpha \pmod{\pi}$. D'autre part la droite passant par les points $B, C$ privée des points $B, C$ est l'ensemble

$$\{M\,;\, M \neq B, C\,;\, (\widehat{\mathbf{MB}, \mathbf{MC}}) \equiv 0 \pmod{\pi}\}.$$

Donc les quatre points $A, B, C, D$ sont alignés si et seulement si $(\widehat{\mathbf{AB}, \mathbf{AC}}) \equiv 0 \pmod{\pi}$ et $(\widehat{\mathbf{DB}, \mathbf{DC}}) \equiv 0 \pmod{\pi}$. D'où le résultat. □

**Définition 7. 7.** *Soit $\alpha \in ]0, \pi[$. L'arc de cercle $\mathcal{C}$ formé des points $M \neq B, C$ tels que*

$$(\widehat{\mathbf{MB}, \mathbf{MC}}) \equiv \alpha \pmod{2\pi} \tag{7.35}$$

*est appelé l'arc capable de la détermination principale $\alpha$ de l'angle orienté* $(\widehat{\widehat{\mathbf{MB}, \mathbf{MC}}})$.

## 7.10 Argument d'un nombre complexe non nul.

Soit $z = x + iy \in \mathbb{C}^*$, $x, y \in \mathbb{R}, x^2 + y^2 \neq 0$. On identifie $z$ au vecteur $\mathbf{u} = \binom{x}{y}$ de l'espace vectoriel euclidien $\mathbb{R}^2$, muni de la base orthonormée canonique $(\mathbf{e_1}, \mathbf{e_2})$ et de l'orientation associée. On a $|z| = \|\mathbf{u}\|$.

**Définition 7. 8.** *On appelle argument de $z$ l'angle orienté* $\widehat{\widehat{(\mathbf{e_1}, \mathbf{u})}}$.

L'argument de $z$ est donc l'élément $A \in SO(2)$ tel que $\mathbf{u} = A\mathbf{e_1}$, *i.e.*

$$\frac{1}{\sqrt{x^2 + y^2}} \begin{pmatrix} x \\ y \end{pmatrix} = A \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{7.36}$$

Dire que $t \in \mathbb{R}$ est une détermination de l'argument de $z$ signifie alors que $A = R(t)$, ou encore :

$$\frac{1}{\sqrt{x^2 + y^2}} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

*i.e.*

$$x + iy = \sqrt{x^2 + y^2}\,(\cos t + i \sin t)$$

ou

$$z = |z| e^{it}.$$

Si $t$ est une détermination de l'argument de $z$, on notera $t = \arg z$. La détermination *principale* de l'argument de $z$ sera notée $\operatorname{Arg} z$.

**Proposition 7. 2.** *Soient $z, w \in \mathbb{C}^*$. Alors*

$$\arg zw \equiv \arg z + \arg w \pmod{2\pi} \tag{7.37}$$

$$\arg \frac{1}{z} \equiv -\arg z \pmod{2\pi} \tag{7.38}$$

$$\arg \frac{z}{w} \equiv \arg z - \arg w \pmod{2\pi} \tag{7.39}$$

**Démonstration.**

Soient $t = \arg z,\ s = \arg w,\ q = \arg(zw)$ des déterminations des arguments. On a $z = |z|e^{it}$ et $w = |w|e^{is}$ donc

$$zw = |z||w|e^{i(t+s)} = |zw|e^{i(t+s)}.$$

Donc $t + s$ est une détermination de l'argument de $zw$, et ainsi

$$q \equiv t + s \pmod{2\pi}$$

d'où (7.37).

Pour (7.38), on a :

$$\frac{1}{z} = \frac{1}{|z|}\, e^{-it} = \left|\frac{1}{z}\right|\, e^{-it},$$

donc $-t$ est une détermination de l'argument de $\frac{1}{z}$.

Enfin, pour (7.39), on a :

$$\begin{aligned} \arg\frac{z}{w} &\equiv \arg z + \arg\frac{1}{w} \pmod{2\pi} \\ &\equiv \arg z - \arg w \pmod{2\pi}. \end{aligned}$$

□

## 7.11 Mesure de l'angle non-orienté de 2 vecteurs.

### 7.11.1 Définition de la mesure de l'angle non-orienté.

On suppose que $E$ est un espace vectoriel euclidien de dimension $n \geqslant 2$. Soient $\mathbf{u}$ et $\mathbf{v}$ deux vecteurs normés non colinéaires et $F$ le plan qu'ils engendrent. D'après le Lemme 7.1 il existe $A \in SO(F)$ unique tel que $f(\mathbf{u}) = \mathbf{v}$. Mais pour pouvoir définir l'angle orienté de $\mathbf{u}$ et $\mathbf{v}$, on a vu qu'il faut fixer une orientation du plan $F$, afin d'identifier canoniquement $SO(F)$ à $SO(2)$. Or si $n \geqslant 3$, la simple donnée d'une orientation de $E$ ne permet pas de définir canoniquement une orientation d'un plan quelconque. On ne peut donc pas définir l'angle orienté de deux vecteurs normés quelconques si $n \geqslant 3$. On définit alors la notion plus faible de *mesure de l'angle non-orienté.*

**Définition 7. 9.** *Soient* $\mathbf{u}, \mathbf{v}$ *deux vecteurs non nuls de l'espace vectoriel euclidien* $E$ *de dimension* $n \geqslant 2$. *On appelle mesure de l'angle non-orienté de* $\mathbf{u}$ *et* $\mathbf{v}$ *l'unique élément* $\theta \in [0, \pi]$ *tel que*

$$\cos\theta = \frac{(\mathbf{u}|\mathbf{v})}{\|\mathbf{u}\|\|\mathbf{v}\|}. \tag{7.40}$$

Cela a bien un sens puisque d'après l'inégalité de Cauchy-Schwarz,

$$-1 \leqslant \frac{(\mathbf{u}|\mathbf{v})}{\|\mathbf{u}\|\|\mathbf{v}\|} \leqslant 1.$$

### 7.11.2 Cas de la dimension 2.

**Proposition 7. 3.** *Soit $E$ un espace vectoriel euclidien orienté de dimension 2, et $\mathcal{B} = (\mathbf{e_1}, \mathbf{e_2})$, une base orthonormée directe de $E$. Soient $\mathbf{u}, \mathbf{v} \in E$ deux vecteurs non nuls, $t$ la détermination principale de l'angle $\widehat{(\mathbf{u}, \mathbf{v})}$, et $\theta$ la mesure de l'angle non-orienté de $\mathbf{u}$ et $\mathbf{v}$. Alors :*
*(i)*

$$|t| = \theta\,;$$

*(ii)*

$$t = \begin{cases} \theta & si\ \det_{\mathcal{B}}(\mathbf{u}, \mathbf{v}) \geqslant 0 \\ -\theta & si\ \det_{\mathcal{B}}(\mathbf{u}, \mathbf{v}) \leqslant 0\,. \end{cases}$$

**Démonstration.**

D'après la parité de la fonction cos , la formule (7.20), et la définition de la mesure de l'angle non-orienté de $\mathbf{u}$ et $\mathbf{v}$, on a

$$\cos|t| = \cos t = \frac{(\mathbf{u}|\mathbf{v})}{\|\mathbf{u}\|\|\mathbf{v}\|} = \cos\theta.$$

Or $|t|$ et $\theta$ appartiennent à $[0, \pi]$, donc $|t| = \theta$.
(ii) Le signe de $t$ est donné par la formule (7.21), d'où le résultat. □

**Théorème 7. 7.** *Soient $B, C$ deux points distincts du plan affine euclidien, et $\alpha \in ]0, \pi[$. L'ensemble des points $M \neq B, C$ tels que la mesure de l'angle non orienté des 2 vecteurs $\mathbf{MB}, \mathbf{MC}$ soit égale à $\alpha$ est constitué de deux arcs de cercle passant par $B, C$, privés des points $B, C$ et symétriques par rapport à la droite passant par $B, C$ (voir Fig. 7.5).*

**Démonstration.**

Choisissons une orientation du plan vectoriel associé. Si $M \neq B, C$, la mesure de l'angle non orienté des 2 vecteurs $\mathbf{MB}, \mathbf{MC}$ est égale à $\alpha$ si et seulement si pour l'angle orienté on a

$$\widehat{(\mathbf{MB}, \mathbf{MC})} \equiv \pm\alpha \qquad (\mathrm{mod}\ 2\pi)\,.$$

Soit $\mathcal{C}$ l'arc capable de la détermination principale $\alpha$ de l'angle orienté $\widehat{(\mathbf{MB}, \mathbf{MC})}$.

Considérons maintenant la *symétrie orthogonale par rapport à la droite passant par $B, C$*. C'est l'application $s$ du plan affine dans lui même définie par

$$s(M) = B + x\mathbf{u} - y\mathbf{v}$$

où $\mathbf{BM} = x\mathbf{u} + y\mathbf{v}$, $\mathcal{R} = (B, (\mathbf{u}, \mathbf{v}))$ étant un repère affine orthonormé direct d'origine $B$, tel que $\mathbf{u}$ soit colinéaire à $\mathbf{BC}$. La partie linéaire $g$ de $s$ est l'endomorphisme du plan vectoriel associé dont la matrice dans la base orthonormée directe $(\mathbf{u}, \mathbf{v})$ est

$$J = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Comme $J \in O(2)$, $g$ est un endomorphisme orthogonal. De ce fait, il conserve le produit scalaire, donc les mesures des angles non orientés. Mais $\det J = -1$, donc

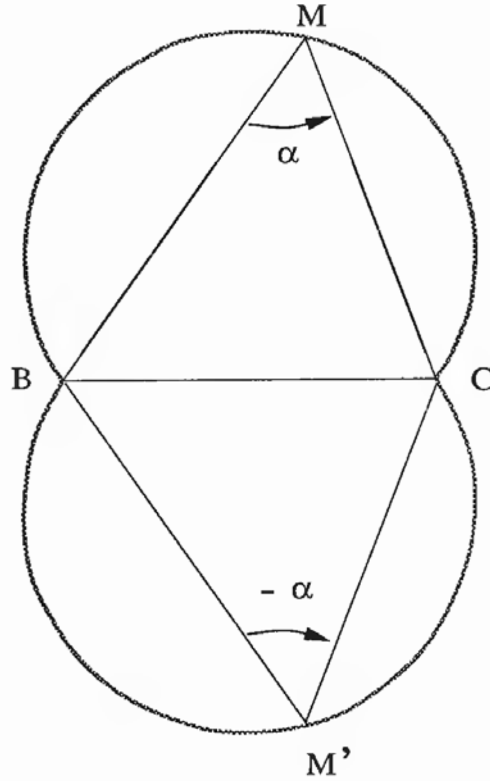



FIG. 7.5: *Arc capable de l'angle non orienté* $\alpha \in ]0, \pi[$.

d'après la formule (7.21), $g$ change le signe des angles orientés. On a donc pour tout point $M$,

$$(\widehat{g(\mathbf{MB}), g(\mathbf{MC})}) \equiv -(\widehat{\mathbf{MB}, \mathbf{MC}}) \pmod{2\pi}. \qquad (7.41)$$

Or comme $s$ est affine de partie linéaire $g$, on a $g(\mathbf{MB}) = \mathbf{s(M)s(B)} = \mathbf{M'B}$ et $g(\mathbf{MC}) = \mathbf{s(M)s(C)} = \mathbf{M'C}$, en notant $M' = s(M)$. (7.41) s'écrit donc :

$$(\widehat{\mathbf{M'B}, \mathbf{M'C}}) \equiv -(\widehat{\mathbf{MB}, \mathbf{MC}}) \pmod{2\pi}.$$

D'où $s(\mathcal{C}) = \{M' \ ; \ M' \neq B, C \ ; \ (\widehat{\mathbf{M'B}, \mathbf{M'C}}) \equiv -\alpha \pmod{2\pi} \ \}$, *i.e.* $s(\mathcal{C})$ est l'arc capable de la détermination principale $-\alpha$ de l'angle orienté $(\widehat{\mathbf{M'B}, \mathbf{M'C}})$. L'ensemble cherché est donc la réunion des 2 arcs symétriques $\mathcal{C}$ et $s(\mathcal{C})$. □

**Corollaire.** *Soient $B, C$ deux points distincts du plan affine euclidien. L'ensemble des points $M \neq B, C$ tels que la mesure de l'angle non orienté des 2 vecteurs* $\mathbf{MB}, \mathbf{MC}$ *soit $\frac{\pi}{2}$ est le cercle de diamètre $BC$, privé des points $B, C$.*

**Démonstration.**

Choisissons une orientation du plan vectoriel associé. Soit $M \neq B, C$. La mesure de l'angle non orienté des 2 vecteurs $\mathbf{MB}, \mathbf{MC}$ est $\frac{\pi}{2}$ si et seulement si pour l'angle orienté on a

$$(\widehat{\mathbf{MB}, \mathbf{MC}}) \equiv \pm\frac{\pi}{2} \pmod{2\pi},$$

ce qui s'écrit encore

$$(\widehat{\mathbf{MB}, \mathbf{MC}}) \equiv \frac{\pi}{2} \pmod{\pi}. \qquad (7.42)$$

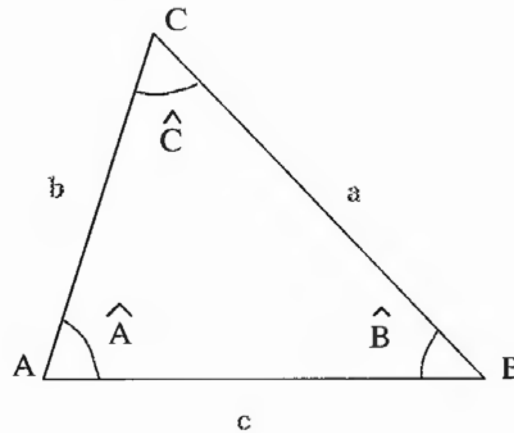



FIG. 7.6: *Relations dans le triangle.*

L'ensemble des points $M \neq B, C$ tels que la mesure de l'angle non orienté des 2 vecteurs $\mathbf{MB}, \mathbf{MC}$ soit $\frac{\pi}{2}$ est donc l'ensemble des points $M \neq B, C$ du plan qui vérifient (7.42). On sait d'après le Th. 7.6 que l'ensemble des points $M \neq B, C$ du plan qui vérifient (7.42) est un cercle $\mathcal{C}$ privé des 2 points $B, C$. Or d'après le Th. 7.7 l'ensemble des points $M \neq B, C$ tels que la mesure de l'angle non orienté des 2 vecteurs $\mathbf{MB}, \mathbf{MC}$ soit $\frac{\pi}{2}$ est constitué de deux arcs de cercle passant par $B, C$, privés des points $B, C$ et symétriques par rapport à la droite passant par $B, C$. Cette réunion n'est un cercle privé de $B, C$ que si $BC$ est un diamètre de chaque arc. □

**Remarque.** (i) On aurait aussi pu dire que le centre du cercle $\mathcal{C}$ est le point $I$ d'intersection de la droite $\mathcal{D}_{\frac{\pi}{2}}$, définie par (7.33) où $\alpha = \frac{\pi}{2}$, avec la médiatrice du segment $[B, C]$. Or $\mathcal{D}_{\frac{\pi}{2}}$ est précisément la droite passant par $B$ et $C$. Donc $I$ est le milieu du segment $[B, C]$.
(ii) On aurait également démontrer directement le Corollaire avec un raisonnement analytique en utilisant un repère affine orthonormé $(I, (\mathbf{u}, \mathbf{v}))$, où $I$ est le milieu du segment $[B, C]$ et $\mathbf{u} = \frac{\mathbf{IC}}{\|\mathbf{IC}\|}$. Si $\|\mathbf{IC}\| = R > 0$, on a dans ce repère : $\mathbf{IB} = \binom{-R}{0}$, $\mathbf{IC} = \binom{R}{0}$, $\mathbf{IM} = \binom{x}{y}$, $\mathbf{MB} = \binom{-R-x}{-y}$, $\mathbf{MC} = \binom{R-x}{-y}$. Les vecteurs $\mathbf{MB}$ et $\mathbf{MC}$ sont orthogonaux si et seulement si $-(R+x)(R-x) + y^2 = 0$, *i.e.* $x^2 + y^2 = R^2$, donc l'ensemble des points $M$ tels que les vecteurs $\mathbf{MB}$ et $\mathbf{MC}$ soient orthogonaux et *non nuls* est le cercle de centre $I$ et de rayon $R$, privé des points $B, C$.

### 7.11.3 Relations métriques dans un triangle.

Soit $ABC$ un triangle du plan affine euclidien. On note $a = \|\mathbf{BC}\|$, $b = \|\mathbf{CA}\|$, $c = \|\mathbf{AB}\|$ et $\widehat{A}, \widehat{B}, \widehat{C}$ désignent les mesures des angles non orientés des vecteurs $\mathbf{AB}$ et $\mathbf{AC}$ , $\mathbf{BC}$ et $\mathbf{BA}$ , $\mathbf{CA}$ et $\mathbf{CB}$ respectivement (voir Fig. 7.6).

**Triangle rectangle.**

Si le triangle $ABC$ est rectangle en $A$, *i.e.* $\widehat{A} = \frac{\pi}{2}$, alors

$$a \cos \widehat{B} = c \tag{7.43}$$

$$a \sin \widehat{B} = b\,. \tag{7.44}$$

En effet, par définition de l'angle non orienté on a $\cos \widehat{B} = \frac{(\mathbf{BC}|\mathbf{BA})}{ac}$. Or

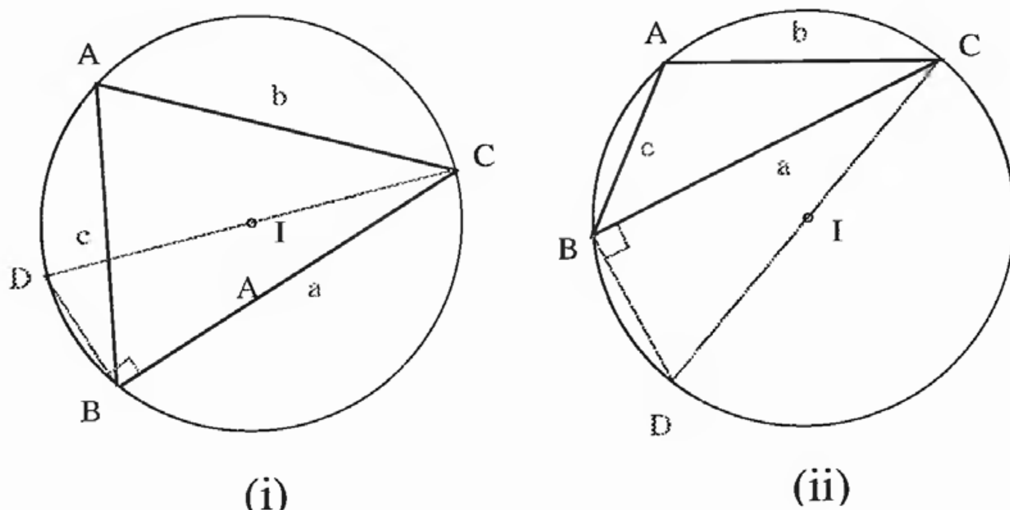



FIG. 7.7: *Loi des sinus.*

$(\mathbf{BC}|\mathbf{BA}) = (\mathbf{AC} - \mathbf{AB}|\mathbf{BA}) = (-\mathbf{AB}|\mathbf{BA}) = c^2$. Donc $\cos \widehat{B} = \frac{c^2}{ac} = \frac{c}{a}$, *i.e.* $a \cos \widehat{B} = c$. L'autre égalité se démontre de la même façon.

**Triangle quelconque.**

$$a^2 = b^2 + c^2 - 2bc \cos \widehat{A}. \tag{7.45}$$

En effet,

$$\begin{aligned} a^2 = \|\mathbf{BC}\|^2 = (\mathbf{BC}|\mathbf{BC}) &= (\mathbf{AC} - \mathbf{AB}|\mathbf{AC} - \mathbf{AB}) \\ &= \|\mathbf{AC}\|^2 + \|\mathbf{AB}\|^2 - 2(\mathbf{AB}|\mathbf{AC}) = b^2 + c^2 - 2bc \cos \widehat{A}. \end{aligned}$$

D'après (7.45) le triangle $ABC$ est rectangle en $A$ si et seulement si

$$a^2 = b^2 + c^2. \tag{7.46}$$

(Cet énoncé constitue le Th. de Pythagore).

**Triangle quelconque : loi des sinus.**

$$\frac{a}{\sin \widehat{A}} = \frac{b}{\sin \widehat{B}} = \frac{c}{\sin \widehat{C}} = 2R \tag{7.47}$$

où $R$ désigne le rayon du cercle circonscrit à $ABC$.

Pour montrer (7.47), notons qu'il y a deux cas possibles suivant que (i) $\widehat{A} \leqslant \frac{\pi}{2}$ ou (ii) $\widehat{A} > \frac{\pi}{2}$ (voir Fig. 7.7). Dans les deux cas (i) et (ii), si $D$ désigne le point diamétralement opposé à $C$ sur le cercle circonscrit, le triangle $CBD$ est rectangle en $B$. Soit $\widehat{D}$ la mesure de l'angle non orienté des vecteurs $\mathbf{DB}$ et $\mathbf{DC}$. Dans le cas (i), $\widehat{D} = \widehat{A}$. Dans le cas (ii), $\widehat{D} = \pi - \widehat{A}$. Donc dans les deux cas, $\sin \widehat{D} = \sin \widehat{A}$. Or on a dans le triangle rectangle $BCD$ :

$$\|\mathbf{DC}\| \sin \widehat{D} = \|\mathbf{BC}\| = a$$

donc $2R\sin\widehat{D} = a$, *i.e.* $2R\sin\widehat{A} = a$. Alors

$$\frac{a}{\sin\widehat{A}} = 2R.$$

En permutant circulairement les rôles de $A, B, C$, on obtiendrait de la même façon $\frac{b}{\sin\widehat{B}} = 2R$ et $\frac{c}{\sin\widehat{C}} = 2R$.

**Triangle quelconque : aire.**

$$2\mathfrak{a} = ab\sin\widehat{C} = bc\sin\widehat{A} = ca\sin\widehat{B} \tag{7.48}$$

où $\mathfrak{a}$ désigne l'aire du triangle $ABC$.

En effet, introduisons un repère affine orthonormé $(O, (\mathbf{e_1}, \mathbf{e_2}))$, et soient $x, y$ les coordonnées associées. Comme dans l'exemple de la sous-section 4.3.3, la surface triangulaire $\Sigma$ définie par $ABC$ peut être paramétrée par $\mathbf{OM} = f(u_1, u_2)$, avec $f : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $f(u_1, u_2) = \mathbf{OA} + u_1\,\mathbf{AB} + u_2\,\mathbf{AC}$, et $(u_1, u_2) \in S = \{(u_1, u_2) \in [0,1] \times [0,1]\ ;\ 0 \leqslant u_1 + u_2 \leqslant 1\ \}$. On a alors

$$\mathfrak{a} = \iint_\Sigma dxdy = \iint_S |J|\, du_1 du_2$$

où

$$J = \det{}_{(\mathbf{e_1},\mathbf{e_2})}(\mathbf{AB}, \mathbf{AC})$$

est le jacobien de $f$. D'après la formule (7.21), on a

$$|J| = \left|\det{}_{(\mathbf{e_1},\mathbf{e_2})}(\mathbf{AB}, \mathbf{AC})\right| = bc\sin\widehat{A}$$

donc

$$\mathfrak{a} = bc\,\sin\widehat{A}\iint_S du_1 du_2\,.$$

Or

$$\iint_S du_1 du_2 = \frac{1}{2}$$

(voir calcul dans l'exemple de la sous-section citée), donc finalement

$$\mathfrak{a} = \frac{1}{2}\,bc\,\sin\widehat{A}\,.$$

En permutant circulairement les rôles de $A, B, C$, on obtiendrait de la même façon

$$\mathfrak{a} = \frac{1}{2}\,ca\,\sin\widehat{B} = \frac{1}{2}\,ab\,\sin\widehat{C}\,.$$

**Remarque.** Si dans la base $(\mathbf{e_1}, \mathbf{e_2})$ on a $\mathbf{AB} = \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$, $\mathbf{AC} = \left(\begin{smallmatrix} x' \\ y' \end{smallmatrix}\right)$, on identifie $\mathbf{AB}$ et $\mathbf{AC}$ aux vecteurs $\left(\begin{smallmatrix} x \\ y \\ 0 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} x' \\ y' \\ 0 \end{smallmatrix}\right)$ de l'espace euclidien $\mathbb{R}^3$. Alors $\mathbf{AB} \wedge \mathbf{AC} = \left(\begin{smallmatrix} x \\ y \\ 0 \end{smallmatrix}\right) \wedge \left(\begin{smallmatrix} x' \\ y' \\ 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 \\ 0 \\ xy'-x'y \end{smallmatrix}\right)$, d'où

$$\left|\det{}_{(\mathbf{e_1},\mathbf{e_2})}(\mathbf{AB}, \mathbf{AC})\right| = |xy' - yx'| = \|\mathbf{AB} \wedge \mathbf{AC}\|.$$

Donc

$$\mathfrak{a} = \frac{1}{2}\,\|\mathbf{AB} \wedge \mathbf{AC}\|.$$

Par permutation circulaire,

$$2\mathfrak{a} = \|\mathbf{AB} \wedge \mathbf{AC}\| = \|\mathbf{BC} \wedge \mathbf{BA}\| = \|\mathbf{CA} \wedge \mathbf{CB}\|\,.$$

## 7.12 Exercices.

**Exercice 7.1.**

Dans le plan euclidien orienté rapporté au repère orthonormé direct $(O, (\mathbf{e_1}, \mathbf{e_2}))$, soient $\mathcal{E}$ l'ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ $(a > b > 0)$ et $F$ le foyer de coordonnées $(c, 0)$ $c > 0$.

(i) Soit $M_0 \in \mathcal{E}$ un point arbitraire de coordonnées $(x_0, y_0)$ et $\mathcal{T}_0$ la tangente à $\mathcal{E}$ au point $M_0$.
(a) Donner un vecteur directeur de la normale à $\mathcal{T}_0$.
(b) Montrer que l'équation de $\mathcal{T}_0$ est $\frac{xx_0}{a^2} + \frac{yy_0}{b^2} = 1$.

(ii) Soit $H$ le projeté orthogonal de $F$ sur $\mathcal{T}_0$. Montrer que les coordonnées $(x, y)$ de $H$ sont solutions du système :

$$\begin{aligned} xx_0b^2 + yy_0a^2 &= a^2b^2 \\ xy_0a^2 - yx_0b^2 &= a^2cy_0 \,. \end{aligned}$$

Calculer $x$ et $y$.

(iii) En élevant au carré chacune des équations du système précédent, et en utilisant le fait que $M_0 \in \mathcal{E}$, montrer que le point $H$ est sur le cercle principal de $\mathcal{E}$. En déduire que le lieu de $H$ quand $M_0$ décrit $\mathcal{E}$ est le cercle principal de $\mathcal{E}$.

(iv) Soient $\mathcal{T}_1$ et $\mathcal{T}_2$ deux tangentes à $\mathcal{E}$ en des points $M_1, M_2$ fixés. Pour $M_0$ différent de $M_1$ et $M_2$ et de leurs symétriques $M_1'$ et $M_2'$ par rapport à $O$, soit $I_1$ (*resp.* $I_2$) le point d'intersection de $\mathcal{T}_0$ avec $\mathcal{T}_1$ (*resp.* $\mathcal{T}_2$), et $P_1$ (*resp.* $P_2$) le projeté orthogonal de $F$ sur $\mathcal{T}_1$ (*resp.* $\mathcal{T}_2$).
(a) Montrer que les points $F, H, I_1, P_1$ sont cocycliques. Montrer qu'il en est de même des points $F, H, I_2, P_2$.
(b) En déduire que

$$(\widehat{\mathbf{FI_1}, \mathbf{FI_2}}) \equiv (\widehat{\mathbf{P_1I_1}, \mathbf{P_2I_2}}) - (\widehat{\mathbf{P_1H}, \mathbf{P_2H}}) \mod \pi \,.$$

(c) Montrer que

$$(\widehat{\mathbf{HP_1}, \mathbf{HP_2}}) \equiv \frac{1}{2}(\widehat{\mathbf{OP_1}, \mathbf{OP_2}}) \mod \pi \,.$$

(d) Qu'en déduit-on?

**Indication.**

(i) (a) Avec le paramétrage habituel $x = a\cos t, y = b\sin t$, le vecteur $\frac{d\mathbf{OM}}{dt} = \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} = \begin{pmatrix} -a\sin t \\ b\cos t \end{pmatrix}$ est un vecteur directeur de la tangente au point de paramètre $t$. Or par dérivation par rapport à $t$ de l'équation de $\mathcal{E}$, on obtient

$$\frac{xx'}{a^2} + \frac{yy'}{b^2} = 0.$$

Cela s'écrit

$$(\mathbf{n} \,|\, \frac{d\mathbf{OM}}{dt}) = 0$$

avec

$$\mathbf{n} = \begin{pmatrix} \frac{x}{a^2} \\ \frac{y}{b^2} \end{pmatrix}.$$

Donc $\mathbf{n}$ est un vecteur directeur de la normale au point de paramètre $t$. $\mathbf{n_0} = \begin{pmatrix} \frac{x_0}{a^2} \\ \frac{y_0}{b^2} \end{pmatrix}$ est un vecteur directeur de la normale à $\mathcal{T}_0$.
(b) L'équation de $\mathcal{T}_0$ est

$$(\mathbf{n_0} \,|\, \mathbf{M_0M}) = 0.$$

Cela s'écrit

$$\frac{x_0(x-x_0)}{a^2}+\frac{y_0(y-y_0)}{b^2}=0$$

ou encore

$$\begin{aligned}\frac{xx_0}{a^2}+\frac{yy_0}{b^2} &= \frac{{x_0}^2}{a^2}+\frac{{y_0}^2}{b^2}\\ &= 1.\end{aligned}$$

(ii) Comme $H \in \mathcal{T}_0$, les coordonnées $x, y$ de $H$ vérifient l'équation de $\mathcal{T}_0$ :

$$xx_0b^2+yy_0a^2=a^2b^2. \tag{7.49}$$

Par ailleurs, le vecteur $\mathbf{FH}$ est colinéaire au vecteur normal $\mathbf{n_0}$ donc

$$\mathbf{FH}\wedge\mathbf{n_0}=0.$$

Comme $\mathbf{FH}=\binom{x-c}{y}$ et $\mathbf{n_0}=\begin{pmatrix}\frac{x_0}{a^2}\\ \frac{y_0}{b^2}\end{pmatrix}$ cela s'écrit

$$xy_0a^2-yx_0b^2=a^2cy_0. \tag{7.50}$$

Maintenant le système des deux équations (7.49), (7.50) a pour déterminant $\Delta=-(b^4x_0^2+a^4y_0^2)<0$. C'est un système de Cramer et la solution est donc

$$x=\frac{\begin{vmatrix}a^2b^2 & y_0a^2\\ a^2cy_0 & -x_0b^2\end{vmatrix}}{\Delta} = \frac{a^2(b^4x_0+a^2cy_0^2)}{b^4x_0^2+a^4y_0^2}, \tag{7.51}$$

$$y=\frac{\begin{vmatrix}x_0b^2 & a^2b^2\\ a^2y_0 & a^2cy_0\end{vmatrix}}{\Delta} = \frac{b^2(-a^2cx_0y_0+a^4y_0)}{b^4x_0^2+a^4y_0^2}. \tag{7.52}$$

(iii) Les équations (7.49) et (7.50) donnent par élévation au carré

$$x^2x_0^2b^4+y^2y_0^2a^4+2xx_0b^2yy_0a^2=a^4b^4$$

$$x^2y_0^2a^4+y^2x_0^2b^4-2xx_0b^2yy_0a^2=c^2a^4y_0^2$$

d'où par addition

$$(x^2+y^2)(x_0^2b^4+a^4y_0^2)=a^4(b^4+c^2y_0^2),$$

donc puisque $c^2=a^2-b^2$ :

$$x^2+y^2=\frac{a^4(b^4+(a^2-b^2)y_0^2)}{x_0^2b^4+a^4y_0^2}.$$

Mais comme le point $M_0$ est sur $\mathcal{E}$, on a :

$$\begin{aligned}x_0^2b^4+a^4y_0^2 &= (a^2-\frac{a^2}{b^2}y_0^2)b^4+a^4y_0^2\\ &= a^2b^4-a^2b^2y_0^2+a^4y_0^2.\end{aligned}$$

Il en résulte que

$$x^2+y^2=a^2$$

donc $H$ est sur le cercle principal de l'ellipse $\mathcal{E}$.

On aurait aussi pu calculer directement $x^2+y^2$ avec les formules (7.51) et (7.52). Cela donne

$$x^2+y^2=a^2\frac{X}{\Delta^2}$$

avec

$$\begin{aligned} X &= a^2b^8x_0^2+a^6c^2y_0^4+a^2b^4c^2x_0^2y_0^2+a^6b^4y_0^2, \\ \Delta^2 &= b^8x_0^4+2a^4b^4x_0^2y_0^2+a^8y_0^4. \end{aligned}$$

Remplaçant $c^2$ par $a^2-b^2$ dans $X$, il vient

$$\begin{aligned} X &= a^2b^8x_0^2+a^8y_0^4-a^6b^2y_0^4+a^4b^4x_0^2y_0^2-a^2b^6x_0^2y_0^2+a^6b^4y_0^2 \\ &= a^8y_0^4+a^4b^4x_0^2y_0^2+a^2b^6x_0^2(b^2-y_0^2)+a^6b^2y_0^2(b^2-y_0^2). \end{aligned}$$

Mais comme $M_0$ est sur $\mathcal{E}$, on a $b^2-y_0^2=\frac{b^2x_0^2}{a^2}$ d'où

$$X=a^8y_0^4+a^4b^4x_0^2y_0^2+b^8x_0^4+a^4b^4x_0^2y_0^2=\Delta^2.$$

Par conséquent

$$x^2+y^2=a^2.$$

Maintenant, quand $M_0$ décrit l'ellipse, on a $x_0=a\cos t, y_0=b\sin t$ et $t$ varie de $0$ à $2\pi$. Les coordonnées $x,y$ du point $H$ sont d'après les formules obtenues des fonctions continues de $t$. Le point $H$ décrit donc une courbe continue, inclue dans le cercle principal. Or pour $t=0$ et $t=2\pi$ on a $x=a, y=0$; pour $t=\pi, x=-a, y=0$. Cette courbe est donc le cercle principal.

(iv) (a) Le triangle $FHI_1$ est rectangle en $H$, donc $H$ est sur le cercle ayant pour diamètre le segment $FI_1$. Le triangle $FP_1I_1$ est rectangle en $P_1$, donc $P_1$ est aussi sur le cercle ayant pour diamètre le segment $FI_1$. Les 4 points $F,H,I_1,P_1$ sont donc cocycliques. On montre de même que les 4 points $F,H,I_2,P_2$ sont cocycliques.

(b) Les 4 points $F,H,I_1,P_1$ étant cocycliques, on a :

$$(\widehat{\mathbf{FI_1},\mathbf{FH}})\equiv(\widehat{\mathbf{P_1I_1},\mathbf{P_1H}}) \pmod{\pi}.$$

De même, les 4 points $F,H,I_2,P_2$ étant cocycliques, on a :

$$(\widehat{\mathbf{FH},\mathbf{FI_2}})\equiv(\widehat{\mathbf{P_2H},\mathbf{P_2I_2}}) \pmod{\pi}.$$

D'où par additivité

$$\begin{aligned} (\widehat{\mathbf{FI_1},\mathbf{FI_2}}) &\equiv (\widehat{\mathbf{P_1I_1},\mathbf{P_1H}})+(\widehat{\mathbf{P_2H},\mathbf{P_2I_2}}) \pmod{\pi} \\ &\equiv (\widehat{\mathbf{P_1I_1},\mathbf{P_2I_2}})-(\widehat{\mathbf{P_1H},\mathbf{P_2H}}) \pmod{\pi}. \end{aligned}$$

(c) $H,P_1,P_2$ sont sur le cercle principal de l'ellipse, donc

$$(\widehat{\mathbf{OP_1},\mathbf{OP_2}})=2(\widehat{\mathbf{HP_1},\mathbf{HP_2}}) \pmod{2\pi}$$

*i.e.*

$$(\widehat{\mathbf{HP_1},\mathbf{HP_2}})=\frac{1}{2}(\widehat{\mathbf{OP_1},\mathbf{OP_2}}) \pmod{\pi}.$$

(d) On en déduit d'après (b) que

$$(\widehat{\mathbf{FI_1}, \mathbf{FI_2}}) \equiv (\widehat{\mathbf{P_1I_1}, \mathbf{P_2I_2}}) - \frac{1}{2}(\widehat{\mathbf{OP_1}, \mathbf{OP_2}}) \qquad (\text{mod } \pi).$$

L'angle $(\widehat{\mathbf{FI_1}, \mathbf{FI_2}})$ a donc une mesure constante modulo $\pi$ lorsque le point $M_0$ varie en restant différent des points $M_1, M_2, M_1', M_2'$ qui sont fixés.

**Exercice 7.2.**

Dans le plan affine euclidien orienté $E$, soient $A, B, C$ 3 points distincts non alignés, $\mathbf{u} = \frac{\mathbf{AB}}{\|\mathbf{AB}\|}$ et $\mathbf{v} = \frac{\mathbf{AC}}{\|\mathbf{AC}\|}$.

(i) Montrer que

$$\Delta = \{A\} \cup \{M \in E\,; M \neq A\,; (\widehat{\mathbf{AB}, \mathbf{AM}}) \equiv (\widehat{\mathbf{AM}, \mathbf{AC}}) \quad (\text{mod } 2\pi)\,\}$$

est une droite affine et que $\mathbf{u} + \mathbf{v}$ est un vecteur directeur de $\Delta$.

(ii) Montrer que

$$\Delta' = \{A\} \cup \{M \in E\,; M \neq A\,; (\widehat{\mathbf{AB}, \mathbf{AM}}) \equiv (\widehat{\mathbf{AM}, -\mathbf{AC}}) \quad (\text{mod } 2\pi)\,\}$$

est une droite affine et que $\mathbf{u} - \mathbf{v}$ est un vecteur directeur de $\Delta'$.

(iii) Montrer que $\Delta$ et $\Delta'$ sont perpendiculaires.

$\Delta$ (*resp.* $\Delta'$) est appelée *bissectrice intérieure (resp. extérieure) de l'angle* $(\widehat{\mathbf{AB}, \mathbf{AC}})$

**Indication.**

(i) On a pour $M \neq A$ :

$$\begin{aligned} M \in \Delta \;&\Leftrightarrow\; (\widehat{\mathbf{AB}, \mathbf{AM}}) \equiv (\widehat{\mathbf{AM}, \mathbf{AC}}) \quad (\text{mod } 2\pi) \\ &\Leftrightarrow\; 2(\widehat{\mathbf{AB}, \mathbf{AM}}) \equiv (\widehat{\mathbf{AB}, \mathbf{AM}}) + (\widehat{\mathbf{AM}, \mathbf{AC}}) \quad (\text{mod } 2\pi) \\ &\Leftrightarrow\; 2(\widehat{\mathbf{AB}, \mathbf{AM}}) \equiv (\widehat{\mathbf{AB}, \mathbf{AC}}) \quad (\text{mod } 2\pi) \end{aligned}$$

ou, en posant $t = (\widehat{\mathbf{AB}, \mathbf{AM}})$, $\theta = (\widehat{\mathbf{AB}, \mathbf{AC}})$ :

$$\begin{aligned} M \in \Delta \;&\Leftrightarrow\; 2t \equiv \theta \quad (\text{mod } 2\pi) \\ &\Leftrightarrow\; t \equiv \frac{\theta}{2} \quad (\text{mod } \pi) \\ &\Leftrightarrow\; \exists k \in \mathbb{Z} \quad t = \frac{\theta}{2} + k\pi \\ &\Leftrightarrow\; \exists k \in \mathbb{Z} \quad \frac{\mathbf{AM}}{\|\mathbf{AM}\|} = R(\frac{\theta}{2} + k\pi)\mathbf{u} \\ &\Leftrightarrow\; \exists k \in \mathbb{Z} \quad \frac{\mathbf{AM}}{\|\mathbf{AM}\|} = (-1)^k R(\frac{\theta}{2})\,\mathbf{u}\,. \end{aligned}$$

Donc $\Delta$ est une droite affine dont un vecteur directeur est $R(\frac{\theta}{2})\,\mathbf{u}$ . Or

$$\begin{aligned} \mathbf{u} + \mathbf{v} &= R(0)\,\mathbf{u} + R(\theta)\,\mathbf{u} \\ &= R(\frac{\theta}{2} - \frac{\theta}{2})\,\mathbf{u} + R(\frac{\theta}{2} + \frac{\theta}{2})\,\mathbf{u} \\ &= R(-\frac{\theta}{2})R(\frac{\theta}{2})\,\mathbf{u} + R(\frac{\theta}{2})R(\frac{\theta}{2})\,\mathbf{u} \\ &= \left(R(-\frac{\theta}{2}) + R(\frac{\theta}{2})\right) R(\frac{\theta}{2})\,\mathbf{u} \\ &= 2\cos\frac{\theta}{2}\, R(\frac{\theta}{2})\,\mathbf{u} \end{aligned}$$

donc $\mathbf{u}+\mathbf{v}$ est un vecteur directeur de $\Delta$ car $\mathbf{u}+\mathbf{v} \neq 0$.
(ii) On a exactement comme en (i) pour $M \neq A$ :

$$\begin{aligned} M \in \Delta' &\Leftrightarrow (\widehat{\mathbf{AB}, \mathbf{AM}}) \equiv (\widehat{\mathbf{AM}, -\mathbf{AC}}) \pmod{2\pi} \\ &\Leftrightarrow 2(\widehat{\mathbf{AB}, \mathbf{AM}}) \equiv (\widehat{\mathbf{AB}, \mathbf{AM}}) + (\widehat{\mathbf{AM}, -\mathbf{AC}}) \pmod{2\pi} \\ &\Leftrightarrow 2(\widehat{\mathbf{AB}, \mathbf{AM}}) \equiv (\widehat{\mathbf{AB}, -\mathbf{AC}}) \pmod{2\pi} \end{aligned}$$

ou, en posant $t = (\widehat{\mathbf{AB}, \mathbf{AM}})$, $\eta = (\widehat{\mathbf{AB}, -\mathbf{AC}})$ :

$$\begin{aligned} M \in \Delta &\Leftrightarrow 2t \equiv \eta \pmod{2\pi} \\ &\Leftrightarrow t \equiv \frac{\eta}{2} \pmod{\pi} \\ &\Leftrightarrow \exists k \in \mathbb{Z} \quad t = \frac{\eta}{2} + k\pi \\ &\Leftrightarrow \exists k \in \mathbb{Z} \quad \frac{\mathbf{AM}}{\|\mathbf{AM}\|} = R(\frac{\eta}{2} + k\pi)\mathbf{u} \\ &\Leftrightarrow \exists k \in \mathbb{Z} \quad \frac{\mathbf{AM}}{\|\mathbf{AM}\|} = (-1)^k R(\frac{\eta}{2})\,\mathbf{u}\,. \end{aligned}$$

Donc $\Delta'$ est une droite affine dont un vecteur directeur est $R(\frac{\eta}{2})\,\mathbf{u}$. Or

$$\begin{aligned} \mathbf{u}-\mathbf{v} &= R(0)\,\mathbf{u} + R(\eta)\,\mathbf{u} \\ &= R(\frac{\eta}{2} - \frac{\eta}{2})\,\mathbf{u} + R(\frac{\eta}{2} + \frac{\eta}{2})\,\mathbf{u} \\ &= R(-\frac{\eta}{2})R(\frac{\eta}{2})\,\mathbf{u} + R(\frac{\eta}{2})R(\frac{\eta}{2})\,\mathbf{u} \\ &= \Big(R(-\frac{\eta}{2}) + R(\frac{\eta}{2})\Big)\, R(\frac{\eta}{2})\,\mathbf{u} \\ &= 2\cos\frac{\eta}{2}\, R(\frac{\eta}{2})\,\mathbf{u} \end{aligned}$$

donc $\mathbf{u}-\mathbf{v}$ est un vecteur directeur de $\Delta'$ car $\mathbf{u}-\mathbf{v} \neq 0$. On notera que

$$(\widehat{\mathbf{AB}, \mathbf{AC}}) \equiv (\widehat{\mathbf{AB}, -\mathbf{AC}}) + (\widehat{-\mathbf{AC}, \mathbf{AC}}) \equiv (\widehat{\mathbf{AB}, -\mathbf{AC}}) + \pi \pmod{2\pi}$$

*i.e.*

$$\theta \equiv \eta \pmod{\pi}.$$

(iii)

$$(\mathbf{u}+\mathbf{v} \,|\, \mathbf{u}-\mathbf{v}) = \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2 = 0\,.$$

**Exercice 7.3.**

Soit $E$ le plan affine euclidien orienté et $(O, (\mathbf{e}_1, \mathbf{e}_2))$ un repère affine orthonormé direct. On appelle *affixe* d'un point $M \in E$ de coordonnées $x, y$, *i.e.* tel que $\mathbf{OM} = x\,\mathbf{e}_1 + y\,\mathbf{e}_2$ (ou d'un vecteur $\mathbf{u} = x\,\mathbf{e}_1 + y\,\mathbf{e}_2$ de composantes $x, y$) le nombre complexe $z = x + iy$.

(i) (a) Soient $A, B$ deux points distincts de $E$, d'affixes respectifs $a, b$. Montrer que $(\widehat{\mathbf{e}_1, \mathbf{AB}})$ est l'argument de $b - a$.
(b) En déduire que si $A, B, C, D$ sont 4 points de $E$ tels que $A \neq B$ et $C \neq D$, ayant pour affixes respectifs $a, b, c, d$ alors

$$(\widehat{\mathbf{AB}, \mathbf{CD}}) \equiv \arg\frac{d-c}{b-a} \pmod{2\pi}.$$

(ii) On appelle *birapport* $(z, z_1, z_2, z_3)$ de 4 nombres complexes $z, z_1, z_2, z_3$ deux-à-deux distincts, et on note $(z, z_1, z_2, z_3)$, le nombre complexe

$$(z, z_1, z_2, z_3) = \frac{z - z_2}{z - z_3} \Big/ \frac{z_1 - z_2}{z_1 - z_3}.$$

Soient $A, B, C, D$ 4 points de $E$ deux-à-deux distincts et $a, b, c, d$ leurs affixes respectifs. Montrer que les points $A, B, C, D$ sont cocycliques ou alignés si et seulement le birapport $(a, b, c, d)$ est réel.

**Indication.**

(i)(a) Notons d'abord que l'affixe du vecteur $\mathbf{AB}$ est $b - a$. En effet, $\mathbf{AB} = \mathbf{OB} - \mathbf{OA} = (b_1 - a_1)\,\mathbf{e_1} + (b_2 - a_2)\,\mathbf{e_2}$ en notant $a = a_1 + ia_2, b = b_1 + ib_2$. Maintenant, $(\widehat{\mathbf{e_1}, \mathbf{AB}})$ est l'unique élément $A \in SO(2)$ tel que $\frac{\mathbf{AB}}{\|\mathbf{AB}\|} = A\mathbf{e_1}$, *i.e.*

$$\frac{1}{\sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2}} \begin{pmatrix} b_1 - a_1 \\ b_2 - a_2 \end{pmatrix} = A \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

En comparant avec la définition (7.36) de l'argument de $z = b - a$, on constate que $A$ est l'argument de $b - a$.

(b) On a d'après (a) :

$$\begin{aligned} (\widehat{\mathbf{e_1}, \mathbf{AB}}) &\equiv \arg(b - a) \qquad (\text{mod } 2\pi) \\ (\widehat{\mathbf{e_1}, \mathbf{CD}}) &\equiv \arg(d - c) \qquad (\text{mod } 2\pi). \end{aligned}$$

Alors

$$\begin{aligned} (\widehat{\mathbf{AB}, \mathbf{CD}}) &\equiv (\widehat{\mathbf{e_1}, \mathbf{CD}}) - (\widehat{\mathbf{e_1}, \mathbf{AB}}) \qquad (\text{mod } 2\pi) \\ &\equiv \arg(d - c) - \arg(b - a) \qquad (\text{mod } 2\pi) \\ &\equiv \arg\frac{d - c}{b - a} \qquad (\text{mod } 2\pi). \end{aligned}$$

(ii) Le birapport $(a, b, c, d)$ appartient à $\mathbb{R}$ si et seulement si

$$\arg(a, b, c, d) \equiv 0 \qquad (\text{mod } \pi).$$

Or par définition du birapport,

$$\begin{aligned} \arg(a, b, c, d) &\equiv \arg\left(\frac{a - c}{a - d} \Big/ \frac{b - c}{b - d}\right) \qquad (\text{mod } 2\pi) \\ &\equiv \arg\frac{a - c}{a - d} - \arg\frac{b - c}{b - d} \qquad (\text{mod } 2\pi) \\ &\equiv (\widehat{\mathbf{DA}, \mathbf{CA}}) - (\widehat{\mathbf{DB}, \mathbf{CB}}) \qquad (\text{mod } 2\pi). \end{aligned}$$

Donc

$$\begin{aligned} (a, b, c, d) \in \mathbb{R} &\iff \arg(a, b, c, d) \equiv 0 \qquad (\text{mod } \pi) \\ &\iff (\widehat{\mathbf{DA}, \mathbf{CA}}) \equiv (\widehat{\mathbf{DB}, \mathbf{CB}}) \qquad (\text{mod } \pi) \\ &\iff (\widehat{\mathbf{AD}, \mathbf{AC}}) \equiv (\widehat{\mathbf{BD}, \mathbf{BC}}) \qquad (\text{mod } \pi) \\ &\iff A, B, C, D \text{ alignés ou cocycliques} \end{aligned}$$

d'après le corollaire du Th. 7.6.

**Exercice 7.4.**

Donner une équation en coordonnées polaires d'un cercle rayon $R > 0$ passant par le pôle.

**Indication.**

On peut prendre $r = 2R\cos(\theta - \varphi)$ avec $(R, \varphi)$ les coordonnées polaires du centre $I$ du cercle.

**Exercice 7.5.**

Quel est le lieu de $w = \frac{z(2-z)}{1-2z}$ lorsque $z$ décrit le cercle unité $\mathbb{T}$ de $\mathbb{C}$?

**Indication.**

Soit $z = e^{it} \in \mathbb{T}$, $t \in \mathbb{R}$. Alors

$$w = \frac{e^{it}(2 - e^{it})}{1 - 2e^{it}} = \frac{2 - e^{it}}{e^{-it} - 2} = -\frac{\zeta}{\bar{\zeta}}$$

en posant $\zeta = 2 - e^{it}$. Donc $|w| = 1$ et le lieu de $w$ est inclus dans le cercle unité $\mathbb{T}$. Maintenant, comme $\arg(-1) \equiv \pi \pmod{2\pi}$ et $\arg\bar{\zeta} \equiv -\arg\zeta \pmod{2\pi}$, on a

$$\arg w \equiv \arg(-1) + \arg\frac{\zeta}{\bar{\zeta}} \equiv \pi + \arg\zeta - \arg\bar{\zeta} \equiv \pi + 2\arg\zeta \qquad \pmod{2\pi}.$$

Or la détermination principale Arg $\zeta$ vérifie (voir Fig.7.8)

$$-\frac{\pi}{6} \leqslant \operatorname{Arg}\zeta \leqslant \frac{\pi}{6}$$

(chaque valeur est prise 2 fois quand $t$ varie de $-\pi$ à $\pi$).

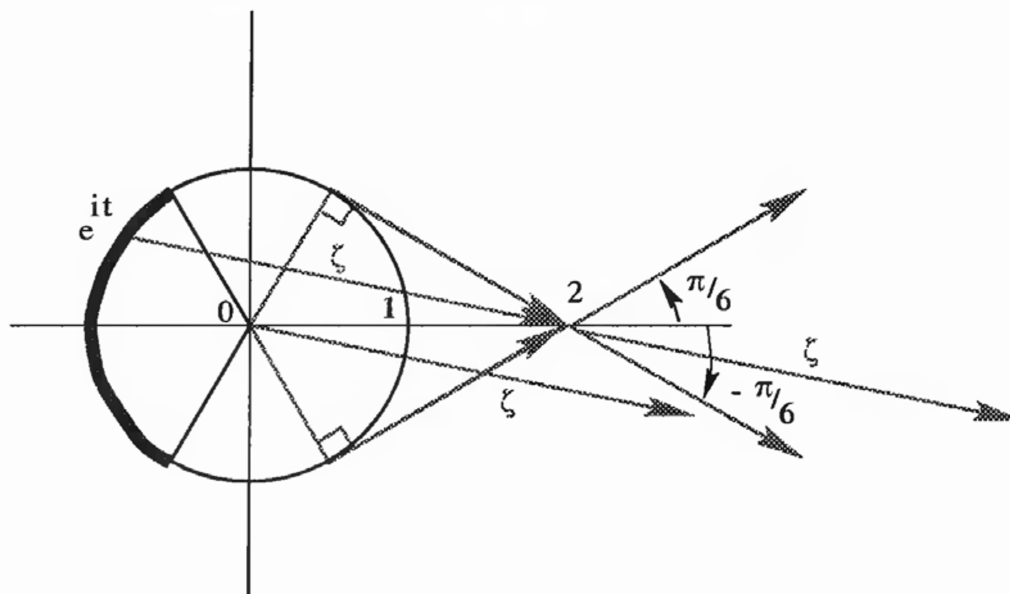



FIG. 7.8: *Exercice 7.5*

Donc

$$\frac{2\pi}{3} \leqslant \pi + 2\operatorname{Arg}\zeta \leqslant \frac{4\pi}{3}.$$

Le lieu cherché est ainsi l'arc $\left\{e^{is} \,;\, \frac{2\pi}{3} \leqslant s \leqslant \frac{4\pi}{3}\right\}$ de $\mathbb{T}$.

# Chapitre 8

# Probabilités.

## 8.1 Lois de probabilité.

### 8.1.1 Expérience aléatoire.

**Définition 8. 1.** *On appelle expérience aléatoire une expérience dont le résultat n'est pas déterminé par les conditions initiales, i.e. dépend du hasard.*

**Exemples.**
(i) Choisir au hasard un entier $\leqslant 100$.
(ii) Lancer un dé (non pipé) et noter le chiffre qui apparaît.
(iii) Lancer 10 fois une pièce de monnaie et compter le nombre de fois où "face" est apparu.
(iv) Compter le nombre de particules émises par un atome radio-actif pendant une période $T$ fixée.
(v) Prendre au hasard un individu d'une population donnée et mesurer sa taille.

### 8.1.2 Modélisation d'une expérience aléatoire.

**Définition 8. 2.** *On appelle ensemble fondamental d'une expérience aléatoire l'ensemble $\Omega$ dont les éléments sont les résultats possibles, ou plus précisément une modélisation des résultats possibles.*

Dans les exemples (i)-(iii), l'ensemble fondamental $\Omega$ est respectivement : $\{1, \cdots, 100\}$, $\{1, 2, 3, 4, 5, 6\}$, $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$. Il est fini. Dans le cas (iv), $\Omega = \mathbb{N}$ est infini dénombrable. Dans ces cas le résultat de l'expérience est un nombre parfaitement déterminé .

Dans le cas (v), on prend comme ensemble fondamental $\Omega = [0, +\infty[$. Mais il n'est pas possible de connaître exactement le résultat de l'expérience : non seulement le résultat d'une mesure est un nombre rationnel, mais en plus ce résultat comporte une erreur due à l'instrument de mesure. La seule issue de la mesure est un *encadrement*, avec une précision que l'on peut choisir arbitrairement, mais qui est finie. Les valeurs exactes ne sont pas observées directement. Par exemple, le résultat de l'expérience pourrait être : "la taille en mètres de l'individu est un élément de l'intervalle $[1.775, 1.785[$", que l'on approchera par la phrase "l'individu mesure 1.78 m.". On est ainsi amené à la notion d'*événement*.

**Définition 8. 3.** *On appelle événement toute partie $A$ de l'ensemble fondamental $\Omega$. Deux événements $A, B$ tels que $A \cap B = \varnothing$ sont dits incompatibles.*

Les événements $\varnothing$ et $\Omega$ sont appelés respectivement événement impossible et événement certain.

La modélisation de l'expérience aléatoire est alors formée de :

1. La donnée d'un *catalogue* d'événement possibles susceptibles d'être réalisés, *i.e.* d'un sous-ensemble $\mathfrak{A}$ de l'ensemble $\mathfrak{P}(\Omega)$ des parties de $\Omega$.

2. La donnée pour chaque $A \in \mathfrak{A}$ d'une *probabilité de réalisation* $P(A) \in [0,1]$, *i.e.* d'une application $P : \mathfrak{A} \to [0,1]$.

On suppose de plus que $\mathfrak{A}$ est une $\sigma$-algèbre et $P$ une mesure de probabilité sur $\mathfrak{A}$ au sens des deux définitions suivantes.

**Définition 8. 4.** *On dit que* $\mathfrak{A} \subset \mathcal{P}(\Omega)$ *est une* $\sigma$*-algèbre si :*

$$\Omega \in \mathfrak{A} \tag{8.1}$$

$$A^c = \complement A \in \mathfrak{A} \qquad \forall A \in \mathfrak{A} \tag{8.2}$$

$$\bigcup_n A_n \in \mathfrak{A} \qquad \forall A_1, A_2, \cdots, A_n, \cdots \in \mathfrak{A} \tag{8.3}$$

Dans la condition (8.3), la famille de parties $A_1, A_2, \cdots, A_n, \cdots \in \mathfrak{A}$ est dénombrable (finie ou infinie). La lettre $\sigma$ se réfère à la fois au terme allemand "Summe" désignant la réunion ensembliste, et au caractère dénombrable de l'union considérée. On notera que si $\mathfrak{A}$ est une $\sigma$-algèbre, on a aussi $\varnothing = \Omega^c \in \mathfrak{A}$. On a également

$$\bigcap_n A_n \in \mathfrak{A} \quad \forall A_1, A_2, \cdots, A_n, \cdots \in \mathfrak{A} \,.$$

En effet,

$$\left(\bigcap_n A_n\right)^c = \bigcup_n (A_n)^c \in \mathfrak{A}$$

d'après (8.2) et (8.3), donc

$$\bigcap_n A_n \in \mathfrak{A}$$

d'après (8.2).

**Définition 8. 5.** *On appelle mesure de probabilité sur la* $\sigma$*-algèbre* $\mathfrak{A}$ *une application* $P : \mathfrak{A} \to [0,1]$ *ayant les 2 propriétés suivantes :*

$$P(\Omega) = 1 \tag{8.4}$$

$$P\left(\bigcup_n A_n\right) = \sum_n P(A_n) \quad \forall A_1, A_2, \cdots, A_n, \cdots \in \mathfrak{A},\ A_i \cap A_j = \varnothing\ \ \forall i \neq j. \tag{8.5}$$

La condition (8.5) s'appelle $\sigma$-additivité . Elle implique en particulier

$$\forall A \in \mathfrak{A}, \quad P(A^c) = 1 - P(A) \tag{8.6}$$

et donc $P(\varnothing) = 0$.

On pose alors :

**Définition 8. 6.** *Une expérience aléatoire est modélisée par un triplet* $(\Omega, \mathfrak{A}, P)$ *où* $\Omega$ *est l'ensemble fondamental,* $\mathfrak{A}$ *une* $\sigma$*-algèbre de parties de* $\Omega$ *et* $P$ *une mesure de probabilité sur* $\mathfrak{A}$*. On dit que le triplet* $(\Omega, \mathfrak{A}, P)$ *est un espace probabilisé ou encore simplement une loi de probabilité sur* $\Omega$*.*

On ne peut pas toujours prendre $\mathfrak{A} = \mathfrak{P}(\Omega)$. Si $\Omega$ est dénombrable (fini ou infini), on peut prendre $\mathfrak{A} = \mathfrak{P}(\Omega)$, mais ce n'est pas toujours le cas si $\Omega$ n'est pas dénombrable. Par exemple, dans le cas de l'exemple (v), les propriétés requises pour $P$, jointes à $\mathfrak{A} = \mathfrak{P}(\Omega)$, mèneraient à une contradiction, donc dans ce cas $\mathfrak{A} \neq \mathfrak{P}(\Omega)$.

Si $\mathfrak{A}$ est une $\sigma$-algèbre, et $\omega \in \Omega$ un élément de l'ensemble fondamental, on appelle *mesure de Dirac au point* $\omega$ et on note $\delta_\omega$ la mesure de probabilité sur $\mathfrak{A}$ définie par

$$\delta_\omega(A) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases}$$

Nous aurons à utiliser plus bas les 2 lemmes suivants.

**Lemme 8. 1.** *Soit $(\Omega, \mathfrak{A}, P)$ un espace probabilisé.*
*(i) Si $A, B \in \mathfrak{A}$ et $A \subset B$, on a $P(A) \leqslant P(B)$.*
*(ii) Si $A_n \in \mathfrak{A}$ et $A_n \subset A_{n+1}$ $\forall n \in \mathbb{N}^*$, on a $P(\bigcup_{n=1}^{+\infty} A_n) = \lim_{n\to+\infty} P(A_n)$.*
*(iii) Si $A_n \in \mathfrak{A}$ et $A_n \supset A_{n+1}$ $\forall n \in \mathbb{N}^*$, on a $P(\bigcap_{n=1}^{+\infty} A_n) = \lim_{n\to+\infty} P(A_n)$.*

**Démonstration.**

(i) $B = A \cup (B \setminus A)$. Comme $A$ et $B \setminus A$ sont disjoints, on a

$$P(B) = P(A) + P(B \setminus A) \geqslant P(A).$$

(ii) La suite $A_n$ est une suite croissante. Posons

$$B_1 = A_1, B_2 = A_2 \setminus A_1, \cdots, B_n = A_n \setminus A_{n-1} \cdots.$$

Alors les $B_k$ sont deux-à-deux disjoints, $A_n = \bigcup_{k=1}^{n} B_k$ et $\bigcup_{n=1}^{+\infty} A_n = \bigcup_{k=1}^{+\infty} B_k$. Or $P(\bigcup_{k=1}^{+\infty} B_k) = \sum_{k=1}^{+\infty} P(B_k)$ est la somme de la série de terme général $P(B_k)$. Par définition de la somme d'une série, on a donc

$$\begin{aligned} P\left(\bigcup_{n=1}^{+\infty} A_n\right) &= P\left(\bigcup_{k=1}^{+\infty} B_k\right) \\ &= \sum_{k=1}^{+\infty} P(B_k) \\ &= \lim_{n\to+\infty} \sum_{k=1}^{n} P(B_k) \\ &= \lim_{n\to+\infty} P(\bigcup_{k=1}^{n} B_k) \\ &= \lim_{n\to+\infty} P(A_n). \end{aligned}$$

(iii) Comme $(A_n)^c \subset (A_{n+1})^c$, on a d'après (ii)

$$P(\bigcup_{n=1}^{+\infty} (A_n)^c) = \lim_{n\to+\infty} P((A_n)^c).$$

Or $\bigcup_{n=1}^{+\infty} (A_n)^c = \left(\bigcap_{n=1}^{+\infty} A_n\right)^c$, donc $1 - P(\bigcap_{n=1}^{+\infty} A_n) = 1 - \lim_{n\to+\infty} P(A_n)$, d'où (iii). □

**Lemme 8. 2.** *Soit $\Omega$ un ensemble. L'intersection $\mathfrak{A} = \bigcap_{i\in I} \mathfrak{A}_i$ d'une famille non vide quelconque $(\mathfrak{A}_i)_{i\in I}$ de $\sigma$-algèbres $\mathfrak{A}_i \subset \mathcal{P}(\Omega)$ est une $\sigma$-algèbre.*

**Démonstration.**

Par hypothèse, $I \neq \emptyset$. On a $\Omega \in \mathfrak{A}_i \ \forall i \in I$, donc $\Omega \in \mathfrak{A}$.

Pour tout $A \in \mathfrak{A}$, on a $A \in \mathfrak{A}_i \ \forall i \in I$, donc $A^c \in \mathfrak{A}_i \ \forall i \in I$, et ainsi $A^c \in \mathfrak{A}$.

Enfin si $A_1, A_2, \cdots, A_n, \cdots \in \mathfrak{A}$, on a $A_1, A_2, \cdots, A_n, \cdots \in \mathfrak{A}_i \ \forall i \in I$, donc $\bigcup_n A_n \in \mathfrak{A}_i \ \forall i \in I$, et ainsi $\bigcup_n A_n \in \mathfrak{A}$. □

**Exemple.** Soit $\mathfrak{S} \subset \mathcal{P}(\Omega)$ un ensemble de parties de $\Omega$. On appelle $\sigma$-algèbre engendrée par $\mathfrak{S}$ la $\sigma$-algèbre intersection de la famille de toutes les $\sigma$-algèbres $\mathfrak{A} \subset \mathcal{P}(\Omega)$ telles que $\mathfrak{S} \subset \mathfrak{A}$. Cette famille n'est pas vide puisqu'elle contient en particulier la $\sigma$-algèbre $\mathcal{P}(\Omega)$. Il est immédiat que la $\sigma$-algèbre engendrée par $\mathfrak{S}$ est la plus petite $\sigma$-algèbre (pour l'inclusion) contenant $\mathfrak{S}$.

## 8.2 Exemples de lois de probabilité.

### 8.2.1 Lois discrètes.

**Equirépartition sur un ensemble fini.**

$\Omega = \{\omega_1, \cdots, \omega_n\}, \mathfrak{A} = \mathfrak{P}(\Omega)$. $P$ est défini par

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\text{NCF}}{\text{NCP}} \quad \forall A \in \mathfrak{A}. \tag{8.7}$$

(NCF signifie nombre de cas favorables et NCP nombre de cas possibles). En particulier

$$P(\{\omega_k\}) = \frac{1}{n} \quad \forall k = 1, \cdots, n.$$

$P$ est bien une loi de probabilité car la $\sigma$-additivité est immédiate et $P(\Omega) = 1$. On peut écrire

$$P = \frac{1}{n} \sum_{k=1}^{n} \delta_{\omega_k}. \tag{8.8}$$

**Loi de Poisson $\mathcal{P}(\lambda)$.**

$\Omega = \mathbb{N}, \mathfrak{A} = \mathfrak{P}(\mathbb{N})$. $\lambda > 0$ est un paramètre, et $P$ est défini par

$$P(A) = \sum_{n \in A} e^{-\lambda} \frac{\lambda^n}{n!} \quad \forall A \in \mathfrak{A}. \tag{8.9}$$

En particulier

$$P(\{n\}) = e^{-\lambda} \frac{\lambda^n}{n!} \quad \forall n \in \mathbb{N}$$

$$P = \sum_{n=0}^{+\infty} e^{-\lambda} \frac{\lambda^n}{n!} \delta_n . \tag{8.10}$$

$P$ est bien une loi de probabilité car la $\sigma$-additivité est immédiate et

$$P(\mathbb{N}) = \sum_{n=0}^{+\infty} P(\{n\}) = e^{-\lambda} \sum_{n=0}^{+\infty} \frac{\lambda^n}{n!} = 1.$$

Cette loi de probabilité est la loi de Poisson de paramètre $\lambda$; on la note $\mathcal{P}(\lambda)$. C'est la loi de probabilité du nombre de particules émises pendant une période $T$ finie par une source radio-active. $\lambda$ est une constante caractéristique du rayonnement.

### 8.2.2 Lois continues.

**Boréliens de $\mathbb{R}$.**

**Définition 8. 7.** *Soit $\Omega = \mathbb{R}$. La $\sigma$-algèbre engendrée par l'ensemble $\mathfrak{I}$ de tous les intervalles (ouverts, fermés ou mixtes) est appelée la $\sigma$-algèbre des boréliens de $\mathbb{R}$. On la note $\mathfrak{B}_{\mathbb{R}}$.*

**Lemme d'unicité.**

**Lemme 8. 3.** *Soient $P_1$ et $P_2$ deux mesures de probabilité sur $\mathfrak{B}_{\mathbb{R}}$. Alors $P_1 = P_2$ si et seulement si $P_1([a,b]) = P_2([a,b]) \quad \forall a, b \in \mathbb{R},\ a < b$.*

**Démonstration.**

La condition est nécessaire. Nous pourrions admettre qu'elle est suffisante, mais nous allons le démontrer en utilisant l'exercice 8.1. L'ensemble $\Lambda$ des parties $A \in \mathfrak{B}_{\mathbb{R}}$ telles que $P_1(A) = P_2(A)$ vérifie les 3 axiomes (a), (b), (c) de l'exercice 8.1 :

- (a) $\Omega \in \Lambda$ puisque $P_1(\Omega) = 1 = P_2(\Omega)$.
- (b) $A \setminus B \in \Lambda$ pour tous $A, B \in \Lambda$ tels que $B \subset A$ puisque $P_1(A \setminus B) = P_1(A) - P_1(B) = P_2(A) - P_2(B) = P_2(A \setminus B)$.
- (c) Si $A_1, A_2, \cdots, A_n, \cdots \in \Lambda$ et $A_n \subset A_{n+1} \quad \forall n \in \mathbb{N}^*$, on a $\bigcup_{n=1}^{+\infty} A_n \in \Lambda$ d'après le Lemme 8.1 (iii) puisque

$$P_1\left(\bigcup_{n=1}^{+\infty} A_n\right) = \lim_{n\to+\infty} P_1(A_n) = \lim_{n\to+\infty} P_2(A_n) = P_2\left(\bigcup_{n=1}^{+\infty} A_n\right).$$

Par hypothèse, $\Lambda$ contient $\{[a,b]\ ,\ a, b \in \mathbb{R},\ a < b\}$. Par union et intersection dénombrables, il est immédiat d'après le Lemme 8.1 (ii) et (iii) que $\Lambda$ contient l'ensemble $\mathfrak{I}$ de tous les intervalles de $\mathbb{R}$. Or $\mathfrak{I}$ vérifie la condition (8.142) de l'exercice cité. D'après le résultat (iv) de cet exercice, $\Lambda$ contient donc la $\sigma$-algèbre $\sigma(\mathfrak{I}) = \mathfrak{B}_{\mathbb{R}}$ engendrée par $\mathfrak{I}$. Or $\Lambda \subset \mathfrak{B}_{\mathbb{R}}$, donc $\Lambda = \mathfrak{B}_{\mathbb{R}}$. Cela prouve que $P_1 = P_2$. □

**Loi définie sur $\mathbb{R}$ par une densité de probabilité.**

**Définition 8. 8.** *Nous appellerons densité de probabilité sur $\mathbb{R}$ toute fonction positive, ayant un nombre fini de points de discontinuité, et telle que l'intégrale $\int_{-\infty}^{+\infty} g(t)\,dt$ soit convergente et égale à 1 :*

$$\int_{-\infty}^{+\infty} g(t)\,dt = 1\ . \tag{8.11}$$

Une densité de probabilité sur $\mathbb{R}$ définit une loi de probabilité sur $\mathfrak{B}_{\mathbb{R}}$ de la façon suivante. Nous admettons que l'application

$$[a,b] \mapsto \int_a^b g(t)\,dt$$

définie sur $\{[a,b]\ ,\ a, b \in \mathbb{R},\ a < b\}$ peut être prolongée en une application $P : \mathfrak{B}_{\mathbb{R}} \to [0,1]$ notée

$$A \in \mathfrak{B}_{\mathbb{R}} \mapsto P(A) = \int_A g(t)\,dt \tag{8.12}$$

vérifiant (8.5). D'après le Lemme 8.3, ce prolongement est unique.

On a bien (8.4) d'après (8.11), car

$$P(\mathbb{R}) = P\left(\bigcup_{n\geqslant 1}[-n,n]\right) = \lim_{n\to+\infty} P([-n,n]) = \lim_{n\to+\infty}\int_{-n}^{n} g(t)\,dt$$
$$= \int_{-\infty}^{+\infty} g(t)\,dt = 1\ .$$

$P$ est donc l'unique mesure de probabilité sur $\mathfrak{B}_{\mathbb{R}}$ telle que

$$P([a,b]) = \int_a^b g(t)\,dt \quad \forall a,b\in\mathbb{R},\ a<b.$$

Pour tous $a,b\in\mathbb{R}$ on a

$$P(\{a\}) = 0\,,$$

$$P([a,b]) = P([a,b[) = P(]a,b[) = P(]a,b]) = \int_a^b g(t)\,dt\,,$$

$$P(]-\infty,a[) = P(]-\infty,a]) = \int_{-\infty}^{a} g(t)\,dt\,.$$

On dira souvent *la loi $P$ sur* $\mathbb{R}$ au lieu de *la loi* $(\mathbb{R},\mathfrak{B}_{\mathbb{R}},P)$.

**Exemples de densités de probabilité sur $\mathbb{R}$.**

**Loi normale $\mathcal{N}(0,1)$.**

$$g(t) = \frac{1}{\sqrt{2\pi}}\,e^{-\frac{t^2}{2}}\,. \tag{8.13}$$

On a bien $\int_{-\infty}^{+\infty} g(t)\,dt = 1$ puisque en posant $t = s\sqrt{2}$,

$$\int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}}\,dt = \sqrt{2}\int_{-\infty}^{+\infty} e^{-s^2}\,ds = \sqrt{2\pi}$$

car on sait que

$$\int_{-\infty}^{+\infty} e^{-s^2}\,ds = \sqrt{\pi}.$$

**Loi normale $\mathcal{N}(m,\sigma)$,** $m\in\mathbb{R}$, $\sigma>0$.

$$g(t) = \frac{1}{\sigma\sqrt{2\pi}}\,e^{-\frac{(t-m)^2}{2\sigma^2}}\,. \tag{8.14}$$

On a bien $\int_{-\infty}^{+\infty} g(t)\,dt = 1$ car le changement de variable $\frac{t-m}{\sigma} = u$ donne:

$$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{(t-m)^2}{2\sigma^2}}\,dt = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}}\,du = 1\,.$$

**Loi uniforme sur $[\alpha,\beta]$, $\alpha,\beta \in \mathbb{R}$.**

$$g(t) = \frac{1}{\beta-\alpha}\, \mathbf{1}_{[\alpha,\beta]}, \tag{8.15}$$

où $\mathbf{1}_{[\alpha,\beta]}(x) = 1$ si $x \in [\alpha,\beta]$ et 0 sinon.

$\int_{-\infty}^{+\infty} g(t)\, dt = 1$ est immédiat.

**Loi exponentielle de paramètre $\lambda > 0$.**

$$g(t) = \lambda e^{-\lambda t}\, \mathbf{1}_{[0,+\infty[}. \tag{8.16}$$

C'est la loi de désintégration d'un atome radio-actif: pour $0 < a < b$, la probabilité que l'atome se désintègre à un instant compris entre $a$ et $b$ est donnée par

$$P([a,b]) = \int_a^b \lambda e^{-\lambda t}\, dt.$$

On a bien

$$\int_{-\infty}^{+\infty} g(t)\, dt = \int_0^{+\infty} \lambda e^{-\lambda t}\, dt = \left[e^{-\lambda t}\right]_0^{+\infty} = 1.$$

## 8.3 Probabilités conditionnelles.

$(\Omega, \mathfrak{A}, P)$ désigne un espace probabilisé.

### 8.3.1 Définition.

**Proposition 8. 1.** *Soit $A \in \mathfrak{A}$ tel que $P(A) > 0$. Posons pour tout $B \in \mathfrak{A}$ :*

$$P(B|A) = \frac{P(B \cap A)}{P(A)}. \tag{8.17}$$

*Alors l'application $P_A : \mathfrak{A} \to [0,1]$ définie par*

$$P_A(B) = P(B|A) \quad \forall B \in \mathfrak{A}$$

*est une mesure de probabilité sur $\Omega$, i.e. $(\Omega, \mathfrak{A}, P_A)$ est une loi de probabilité.*

**Démonstration.**

Il suffit de vérifier que l'application $P_A : \mathfrak{A} \to [0,1]$ possède les 2 propriétés (8.4) et (8.5). Pour (8.4),

$$P_A(\Omega) = \frac{P(\Omega \cap A)}{P(A)} = \frac{P(A)}{P(A)} = 1\,.$$

Pour (8.5), soient $A_1, A_2, \cdots, A_n, \cdots \in \mathfrak{A}$ tels que $A_i \cap A_j = \varnothing\ \forall i \neq j$ . Alors

$$\begin{aligned} P_A\left(\bigcup_n A_n\right) &= \frac{P\left(A \cap \left(\bigcup_n A_n\right)\right)}{P(A)} \\ &= \frac{P\left(\bigcup_n (A \cap A_n)\right)}{P(A)} \\ &= \frac{\sum_n P\left(A \cap A_n\right)}{P(A)} \\ &= \sum_n \frac{P\left(A \cap A_n\right)}{P(A)} \\ &= \sum_n P_A(A_n)\,. \end{aligned}$$

□

**Définition 8. 9.** *La loi de probabilité* $(\Omega, \mathfrak{A}, P_A)$ *s'appelle la loi de probabilité conditionnelle sur* $\Omega$ *associée à l'événement* $A \in \mathfrak{A}$.

La définition (8.17) s'écrit encore

$$P(A \cap B) = P(B|A)P(A) \quad \forall A, B \in \mathfrak{A}. \tag{8.18}$$

### 8.3.2 Événements indépendants.

**Définition 8. 10.** *Deux événements* $A, B \in \mathfrak{A}$ *sont dits indépendants si*

$$P(A \cap B) = P(A)P(B). \tag{8.19}$$

D'après la formule (8.18), les événements $A$ et $B$ sont indépendants si et seulement si $P(B|A) = P(B)$.

**Définition 8. 11.** *Une famille finie ou infinie* $(A_i)_{i\in I}$ *d'événements est dite indépendante (on dit encore que les événements* $A_i$, $i \in I$ *sont indépendants dans leur ensemble) si pour toute famille finie* $(A_{i_1}, A_{i_2} \cdots, A_{i_n})$ *extraite de la famille* $(A_i)_{i\in I}$ *on a*

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_n}) = P(A_{i_1})P(A_{i_2})\ldots P(A_{i_n}). \tag{8.20}$$

On notera qu'une famille formée d'événements deux-à-deux indépendants *n'est pas* nécessairement indépendante.

En effet, considérons l'expérience consistant à jeter simultanément 2 pièces (équilibrées) de monnaie (marquées 1 et 2) et à observer les couples de résultats obtenus. On a $\Omega = \{(P,P), (P,F), (F,P), (F,F)\}$, $P$ désignant "pile" et $F$ désignant "face", et le 1er terme du couple référant à la pièce 1, le 2ème à la pièce 2. Il est clair que les 4 éléments de $\Omega$ sont équiprobables. Soient $A$ l'événement "la pièce 1 donne face", $B$ l'événement "la pièce 2 donne pile", et $C$ l'événement "les pièces 1 et 2 donnent toutes les deux pile ou toutes les deux face". On a

$$A = \{(F,P), (F,F)\},\ B = \{(F,P), (P,P)\},\ C = \{(F,F), (P,P)\}.$$

Donc $P(A) = P(B) = P(C) = \frac{1}{2}$. Maintenant

$$A \cap B = \{(F,P)\},\ A \cap C = \{(F,F)\},\ B \cap C = \{(P,P)\},$$

donc $P(A \cap B) = \frac{1}{4} = P(A)P(B)$, $P(A \cap C) = \frac{1}{4} = P(A)P(C)$, $P(B \cap C) = \frac{1}{4} = P(B)P(C)$. Les événements $A, B, C$ sont deux-à-deux indépendants. Mais la famille $\{A, B, C\}$ n'est pas indépendante puisque $P(A \cap B \cap C) = 0$.

### 8.3.3 Formule de Bayes.

On appellera *système complet d'événements* une *partition* dénombrable (finie ou infinie) $\Omega = \bigcup_n A_n$ avec $A_n \in \mathfrak{A}$, $P(A_n) \neq 0\ \forall n$. On a $A_n \cap A_m = \varnothing$ pour tous $n, m, n \neq m$ par définition d'une partition.

**Théorème 8. 1 (Bayes).** *Soit* $(A_n)$ *un système complet d'événements. Pour tout événement* $B \in \mathfrak{A}$ *tel que* $P(B) \neq 0$, *on a :*

$$P(A_k|B) = \frac{P(B|A_k)P(A_k)}{\sum_n P(B|A_n)P(A_n)} \quad \forall k. \tag{8.21}$$

**Démonstration.**

La formule des probabilités conditionnelles (8.17) donne

$$P(B \cap A_k) = P(A_k|B)P(B) = P(B|A_k)P(A_k).$$

Or comme $\Omega$ est réunion disjointe des $A_n$, on a

$$P(B) = P\left(B \cap \bigcup_n A_n\right) = \sum_n P(B \cap A_n) = \sum_n P(B|A_n)P(A_n),$$

d'où (8.21). □

# 8.4 Variables aléatoires réelles.

## 8.4.1 Notion de variable aléatoire réelle.

**Variable aléatoire.**

**Définition 8. 12.** *Soit $(\Omega, \mathfrak{A}, P)$ un espace probabilisé. On appelle variable aléatoire réelle sur $(\Omega, \mathfrak{A}, P)$ ou plus simplement sur $\Omega$, une application $X : \Omega \to \mathbb{R}$ ayant la propriété suivante :*

$$X^{-1}(]-\infty, a[) = \{\omega \in \Omega;\ X(\omega) \in ]-\infty, a[\} \in \mathfrak{A} \quad \forall a \in \mathbb{R}. \tag{8.22}$$

*La variable aléatoire $X$ est dite discrète si l'ensemble des valeurs prises par $X$ est dénombrable (fini ou infini); elle est dite continue si l'ensemble des valeurs prises par $X$ est un intervalle non réduit à un point.*

Nous abrégerons variable aléatoire réelle en v.a.

**Lemme 8. 4.** *La condition (8.22) implique*

$$X^{-1}(B) = \{\omega \in \Omega;\ X(\omega) \in B\} \in \mathfrak{A} \quad \forall B \in \mathfrak{B}_{\mathbb{R}} \tag{8.23}$$

*où $\mathfrak{B}_{\mathbb{R}}$ est la $\sigma$-algèbre des boréliens de $\mathbb{R}$.*

**Démonstration.**

Montrons d'abord que l'ensemble de parties

$$\mathfrak{C} = \{B \in \mathfrak{B}_{\mathbb{R}}\,;\, X^{-1}(B) \in \mathfrak{A}\} \subset \mathfrak{B}_{\mathbb{R}}$$

est une $\sigma$-algèbre de parties de $\mathbb{R}$. Si $B \in \mathfrak{C}$, $B \in \mathfrak{B}_{\mathbb{R}}$ implique $B^c \in \mathfrak{B}_{\mathbb{R}}$ et d'autre part $X^{-1}(B) \in \mathfrak{A}$ implique

$$X^{-1}(B^c) = (X^{-1}(B))^c \in \mathfrak{A}.$$

Donc $B^c \in \mathfrak{C}$, et ainsi $\mathfrak{C}$ est stable par passage au complémentaire. Ensuite, soient $B_1, B_2, \cdots, B_n, \cdots \in \mathfrak{C}$. Comme $B_n \in \mathfrak{B}_{\mathbb{R}}$ pour tout $n$, on a $\bigcup_n B_n \in \mathfrak{B}_{\mathbb{R}}$. D'autre part

$$X^{-1}\left(\bigcup_n B_n\right) = \bigcup_n X^{-1}(B_n) \in \mathfrak{A}$$

puisque $X^{-1}(B_n) \in \mathfrak{A}$ pour tout $n$. Ainsi $\mathfrak{C}$ est stable par union dénombrable. Enfin, si l'on pose $B_n = ]-\infty, n[\ \forall\, n \in \mathbb{N}^*$, on a d'après (8.22), $X^{-1}(B_n) \in \mathfrak{A}$ donc $B_n \in \mathfrak{C}\ \forall\, n \in \mathbb{N}^*$ et alors $\mathbb{R} = \bigcup_n B_n \in \mathfrak{C}$ puisque $\mathfrak{C}$ est stable par union dénombrable. On a donc bien obtenu que $\mathfrak{C}$ est une $\sigma$-algèbre de parties de $\mathbb{R}$.

Maintenant, d'après (8.22), $\mathfrak{C}$ contient tout intervalle de la forme $]-\infty, a[,\ \forall a \in \mathbb{R}$. Comme $[x,y[= ]-\infty, y[ \setminus ]-\infty, x[$ , $(x, y \in \mathbb{R}, x < y)$, $\mathfrak{C}$ contient tous les intervalles de $\mathbb{R}$ de cette forme. Par union et intersection dénombrables, il est immédiat que $\mathfrak{C}$ contient tous les intervalles de $\mathbb{R}$. On en déduit que $\mathfrak{C}$ contient la $\sigma$-algèbre $\mathfrak{B}_{\mathbb{R}}$ des boréliens. Or $\mathfrak{C} \subset \mathfrak{B}_{\mathbb{R}}$, donc $\mathfrak{C} = \mathfrak{B}_{\mathbb{R}}$. □

**Exemple.** Soit $(\Omega, \mathfrak{A}, P)$ un espace probabilisé et $X : \Omega \to \mathbb{R}$ une application dont l'ensemble des valeurs est la famille $(x_i)_{i\in I}$ avec $I$ dénombrable (fini ou infini). Alors $X$ est une v.a. (discrète) si et seulement si la condition suivante est vérifiée :

$$X^{-1}(\{x_i\}) = \{\omega \in \Omega;\ X(\omega) = x_i\} \in \mathfrak{A} \quad \forall i \in I. \tag{8.24}$$

En effet, la condition (8.24) est nécessaire d'après le Lemme 8.4; elle est suffisante puisque si elle est vérifiée on a pour tout $a \in \mathbb{R}$

$$X^{-1}(]-\infty, a[) = \bigcup_{x_i < a} X^{-1}(\{x_i\}) \in \mathfrak{A}.$$

**Opérations algébriques sur les variables aléatoires réelles.**

**Lemme 8. 5.** *Soient $X, Y$ deux variables aléatoires réelles sur un espace probabilisé $(\Omega, \mathfrak{A}, P)$. Alors $X+Y$, $\lambda X$ $(\lambda \in \mathbb{R})$, $|X|$, $XY$, et $\frac{1}{X}$ si $X$ ne s'annule pas, sont des variables aléatoires réelles. Une application constante est une v.a.*

**Démonstration.**

• Cas d'une constante. Si $X$ est constante, il existe $c \in \mathbb{R}$ telle que $X(\omega) = c\ \forall \omega \in \Omega$. Alors $X^{-1}(]-\infty, a[) = \varnothing$ ou $\mathbb{R}$ suivant que $a \leqslant c$ ou $a > c$. Donc (8.22) est vérifiée.

• Cas de $X+Y$. Soit $a \in \mathbb{R}$. Pour tout $\omega \in \Omega$, on a

$$\omega \in (X+Y)^{-1}(]-\infty, a[) \Leftrightarrow X(\omega) + Y(\omega) < a$$

$$\Leftrightarrow\ X(\omega) < a - Y(\omega)$$
$$\Leftrightarrow\ \exists\, r \in \mathbb{Q}\quad X(\omega) < r < a - Y(\omega)$$

$$\Leftrightarrow\ \exists\, r \in \mathbb{Q}\quad \omega \in X^{-1}(]-\infty, r[) \cap Y^{-1}(]-\infty, a-r[)$$
$$\Leftrightarrow\ \omega \in \bigcup_{r\in\mathbb{Q}} \left(X^{-1}(]-\infty, r[) \cap Y^{-1}(]-\infty, a-r[)\right).$$

Donc

$$(X+Y)^{-1}(]-\infty, a[) = \bigcup_{r\in\mathbb{Q}} \left(X^{-1}(]-\infty, r[) \cap Y^{-1}(]-\infty, a-r[)\right) \in \mathfrak{A}$$

puisque $X$ et $Y$ sont des v.a., *i.e.* vérifient la condition (8.22) et $\mathbb{Q}$ est dénombrable. $a \in \mathbb{R}$ étant arbitraire, on a alors bien obtenu la condition (8.22) pour $X+Y$, donc

$X+Y$ est une v.a.
• Cas de $\lambda X$. Si $\lambda = 0$, $\lambda X$ est la constante 0, donc c'est une v.a. Si $\lambda > 0$,

$$(\lambda X)^{-1}(]-\infty, a[) = X^{-1}(]-\infty, \frac{a}{\lambda}[) \in \mathfrak{A} \quad \forall\, a \in \mathbb{R}$$

donc (8.22) est vérifiée. Si $\lambda < 0$,

$$(\lambda X)^{-1}(]-\infty, a[) = X^{-1}(]\frac{a}{\lambda}, +\infty[) \in \mathfrak{A} \quad \forall\, a \in \mathbb{R}$$

d'après le Lemme 8.4, donc (8.22) est vérifiée.
• Cas de $|X|$. Si $a \leqslant 0$, $|X|^{-1}(]-\infty, a[) = \varnothing \in \mathfrak{A}$. Si $a > 0$, $|X|^{-1}(]-\infty, a[) = |X|^{-1}([0, a[) = X^{-1}(]-a, a[) \in \mathfrak{A}$ d'après le Lemme 8.4. Donc $|X|$ est une v.a.
• Cas de $XY$. D'abord $X^2$ est une v.a. En effet, si $a \leqslant 0$, $(X^2)^{-1}(]-\infty, a[) = \varnothing \in \mathfrak{A}$. Si $a > 0$, $(X^2)^{-1}(]-\infty, a[) = (X^2)^{-1}([0, a[) = (|X|)^{-1}(]-\sqrt{a}, \sqrt{a}[) \in \mathfrak{A}$ d'après le Lemme 8.4, appliqué à la v.a. $|X|$. Donc $(X^2)^{-1}(]-\infty, a[) \in \mathfrak{A}$ et $X^2$ est une v.a.

Maintenant,

$$XY = \frac{1}{2}\left((X+Y)^2 - X^2 - Y^2\right).$$

Or $X+Y, (X+Y)^2, -X^2, -Y^2$ sont des v.a. donc aussi $XY$.
• Cas de $\frac{1}{X}$. Soit $a > 0$. Pour $\omega \in \Omega$, on a $\frac{1}{X(\omega)} < a$ si et seulement si $X(\omega) < 0$ ou $X(\omega) > \frac{1}{a}$. Donc

$$\left(\frac{1}{X}\right)^{-1}(]-\infty, a[) = X^{-1}(]-\infty, 0[\cup]\frac{1}{a}, +\infty[) \in \mathfrak{A}$$

d'après le Lemme 8.4.
Soit $a < 0$. Pour $\omega \in \Omega$, on a $\frac{1}{X(\omega)} < a$ si et seulement si $\frac{1}{a} < X(\omega) < 0$. Donc

$$\left(\frac{1}{X}\right)^{-1}(]-\infty, a[) = X^{-1}(]\frac{1}{a}, 0[) \in \mathfrak{A}$$

d'après le Lemme 8.4.
Enfin pour $a = 0$, $(\frac{1}{X})^{-1}(]-\infty, 0[) = X^{-1}(]-\infty, 0[) \in \mathfrak{A}$. □

**Corollaire.** *L'ensemble des v.a. réelles sur un espace probabilisé* $(\Omega, \mathfrak{A}, P)$ *est un sous-espace vectoriel de l'espace vectoriel des applications de* $\Omega$ *dans* $\mathbb{R}$ *contenant les constantes.*

**Démonstration.**

D'après lemme, $X+Y$ et $\lambda X$ ($\lambda \in \mathbb{R}$) sont des v.a. D'autre part les constantes sont des v.a. d'après le lemme. □

**Loi de probabilité d'une variable aléatoire.**

Soit $(\Omega, \mathfrak{A}, P)$ un espace probabilisé, et $X$ une variable aléatoire réelle sur $\Omega$. D'après (8.23), $X^{-1}(B) \in \mathfrak{A}$ pour tout borélien $B \in \mathfrak{B}_{\mathbb{R}}$, de sorte que $P(X^{-1}(B))$ est défini pour tout $B \in \mathfrak{B}_{\mathbb{R}}$. On définit alors une loi de probabilité $P_X$ sur $\mathbb{R}$ en posant

$$P_X(B) = P(X^{-1}(B)) \quad \forall B \in \mathfrak{B}_{\mathbb{R}}. \tag{8.25}$$

On dit que $P_X$ est la *loi de probabilité de* $X$, ou encore que $X$ suit la loi $P_X$. $P_X$ est bien une loi de probabilité sur $\mathbb{R}$. En effet, $P_X(\mathbb{R}) = P(X^{-1}(\mathbb{R})) = P(\Omega) = 1$, et si $(B_n)_{n\in\mathbb{N}}$ est une suite de boréliens deux-à-deux disjoints, on a :

$$\begin{aligned} P_X\left(\bigcup_n B_n\right) &= P\left(X^{-1}\left(\bigcup_n B_n\right)\right) \\ &= P\left(\bigcup_n X^{-1}(B_n)\right) = \sum_n P\left(X^{-1}(B_n)\right) = \sum_n P_X(B_n) \end{aligned}$$

en notant que les $X^{-1}(B_n)$ sont deux-à-deux disjoints.

On notera que si $P$ est une mesure de probabilité quelconque sur $\mathfrak{B}_{\mathbb{R}}$, il existe une v.a. $X$ sur l'espace probabilisé $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}}, P)$ telle que $P = P_X$ : il suffit de prendre pour application $X : \mathbb{R} \to \mathbb{R}$ l'application identité $X(\omega) = \omega \ \ \forall \omega \in \mathbb{R}$.

**Motivations.**

La notion de variable aléatoire permet entre autres :
▶ d'introduire des lois de probabilités correspondant à des situations très concrètes;
▶ de modéliser dans certains cas une expérience aléatoire de façon à se ramener à un espace probabilisé *avec équirépartition*, le résultat numérique de l'expérience apparaissant alors comme une v.a. sur cet espace;
▶ de modéliser certaines situations dépendant à la fois du hasard et du temps par une ou des v.a. sur un espace probabilisé.

**Notation.**

Si $X$ est une v.a. sur $(\Omega, \mathfrak{A}, P)$, nous noterons : $P(X \in B) = P_X(B)$ pour $B \in \mathfrak{B}_{\mathbb{R}}$, $P(X = t) = P_X(\{t\})$ pour $t \in \mathbb{R}$, $P(a \leqslant X \leqslant b) = P_X([a, b])$, etc.

**Fonction de répartition d'une v.a.**

**Définition 8. 13.** *Soit $X$ une v.a. sur un espace probabilisé $(\Omega, \mathfrak{A}, P)$. On appelle fonction de répartition de $X$ l'application $F_X : \mathbb{R} \to \mathbb{R}$ définie par*

$$F_X(x) = P_X(]-\infty, x]) = P(X \leqslant x). \tag{8.26}$$

**Lemme 8. 6.** *Soit $X$ une v.a. sur $(\Omega, \mathfrak{A}, P)$. Sa fonction de répartition $F_X : \mathbb{R} \to \mathbb{R}$ a les propriétés suivantes :*
*(i) $F_X$ est croissante,*
*(ii) $F_X$ est continue à droite,*
*(iii)* $\lim_{x\to-\infty} F_X(x) = 0$, $\lim_{x\to+\infty} F_X(x) = 1$.

**Démonstration.**

(i) Pour $x \leqslant y$, on a $]-\infty, x] \subset ]-\infty, y]$ donc

$$F_X(x) = P_X(]-\infty, x]) \leqslant P_X(]-\infty, y]) = F_X(y)$$

et $F_X$ est croissante.
(ii) Soit $a \in \mathbb{R}$ et $(x_n)$ une suite décroissante telle que $a = \lim_{n\to+\infty} x_n$. La suite d'intervalles $]-\infty, x_n]$ est telle que $]-\infty, x_n] \supset ]-\infty, x_{n+1}]$ et $]-\infty, a] = \bigcap_{n=1}^{\infty} ]-\infty, x_n]$. Donc d'après le Lemme 8.1(iii),

$$F_X(a) = P_X(]-\infty, a]) = \lim_{n\to+\infty} P_X(]-\infty, x_n]) = \lim_{n\to+\infty} F_X(x_n).$$

On en déduit que $F_X$ est continue à droite au point $a$.
(iii) De même, soit $(x_n)$ une suite décroissante tendant vers $-\infty$. La suite d'intervalles $]-\infty, x_n]$ est telle que $]-\infty, x_n] \supset ]-\infty, x_{n+1}]$ et

$$\bigcap_{n=1}^{\infty} ]-\infty, x_n] = \varnothing.$$

Donc

$$\lim_{n\to+\infty} F_X(x_n) = \lim_{n\to+\infty} P_X(]-\infty, x_n]) = P_X\left(\bigcap_{n=1}^{\infty} ]-\infty, x_n]\right) = P_X(\varnothing) = 0.$$

On en déduit que $\lim_{x\to-\infty} F_X(x) = 0$.
Enfin, soit $(x_n)$ une suite croissante tendant vers $+\infty$. La suite d'intervalles $]-\infty, x_n]$ est telle que $]-\infty, x_n] \subset ]-\infty, x_{n+1}]$ et $\bigcup_{n=1}^{\infty} ]-\infty, x_n] = \mathbb{R}$. Donc d'après le Lemme 8.1(ii),

$$\lim_{n\to+\infty} F_X(x_n) = \lim_{n\to+\infty} P_X(]-\infty, x_n]) = P_X\left(\bigcup_{n=1}^{\infty} ]-\infty, x_n]\right) = P_X(\mathbb{R}) = 1.$$

On en déduit que $\lim_{x\to+\infty} F_X(x) = 1$. □

**Lemme 8. 7.** *Soient $X, Y$ deux v.a. sur $(\Omega, \mathfrak{A}, P)$, $P_X, P_Y$ leurs lois de probabilité et $F_X, F_Y$ leurs fonctions de répartition. On a $P_X = P_Y$ si et seulement si $F_X = F_Y$.*

**Démonstration.**
Il est clair que $P_X = P_Y$ implique $F_X = F_Y$. Réciproquement, supposons $F_X = F_Y$. Alors $P_X(]-\infty, x]) = P_Y(]-\infty, x])\ \forall x \in \mathbb{R}$. Soit $\mathfrak{S} = \{]-\infty, x]\,;\, x \in \mathbb{R}\}$. Comme dans la démonstration du Lemme 8.3, l'ensemble $\Lambda$ des parties $A \in \mathfrak{B}_{\mathbb{R}}$ telles que $P_X(A) = P_Y(A)$ vérifie les 3 axiomes (a), (b), (c) de l'exercice 8.1. Or $\mathfrak{S} \subset \Lambda$ et $\mathfrak{S}$ vérifie la condition (8.142) de l'exercice cité. D'après le résultat (iv) de cet exercice, $\Lambda$ contient donc la $\sigma$-algèbre $\sigma(\mathfrak{S})$ engendrée par $\mathfrak{S}$.

Mais comme $]x, y] = ]-\infty, y] \setminus ]-\infty, x]$, $(x, y \in \mathbb{R}, x \leqslant y)$, $\sigma(\mathfrak{S})$ contient tous les intervalles de $\mathbb{R}$ de cette forme. Par union et intersection dénombrables, il est immédiat que $\sigma(\mathfrak{S})$ contient tous les intervalles de $\mathbb{R}$. On en déduit que $\sigma(\mathfrak{S})$ contient la $\sigma$-algèbre $\mathfrak{B}_{\mathbb{R}}$ des boréliens et donc est égal à $\mathfrak{B}_{\mathbb{R}}$. Alors $\Lambda$ contient $\mathfrak{B}_{\mathbb{R}}$. Or $\Lambda \subset \mathfrak{B}_{\mathbb{R}}$, d'où l'égalité. Cela prouve que $P_X = P_Y$. □

**Exemple.** Soit $X$ une v.a. suivant la loi normale centrée réduite $\mathcal{N}(0,1)$. La fonction de répartition de $X$ est

$$\Psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt\,. \tag{8.27}$$

Posons pour tout $x > 0$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{0}^{x} e^{-\frac{t^2}{2}}\, dt\,. \tag{8.28}$$

Les valeurs de $\Phi$ sont données par la Table 1. On a alors pour tout $x > 0$

$$P(-x \leqslant X \leqslant x) = 2\Phi(x) \tag{8.29}$$

$$\Psi(x) = P(X \leqslant x) = \frac{1}{2} + \Phi(x). \tag{8.30}$$

**Variables aléatoires indépendantes.**

**Définition 8. 14.** *Deux v.a. $X, Y$ sur $(\Omega, \mathfrak{A}, P)$ sont dites indépendantes si pour tous $A, B \in \mathfrak{B}_{\mathbb{R}}$ les événements $\{X \in A\}$ et $\{Y \in B\}$ de $\mathfrak{A}$ sont indépendants, i.e. si*

$$P(X \in A\,,\, Y \in B) = P(X \in A)\, P(Y \in B) \tag{8.31}$$

*en notant*

$$P(X \in A\,,\, Y \in B) = P(\{X \in A\} \cap \{Y \in B\}).$$

*Une famille finie ou infinie $(X_i)_{i\in I}$ de v.a. sur $(\Omega, \mathfrak{A}, P)$ est dite indépendante (on dit encore que les v.a. $X_i$, $i \in I$ sont indépendantes dans leur ensemble) si pour toute famille finie $(X_{i_1}, X_{i_2}, \cdots, X_{i_n})$ extraite de la famille $(X_i)_{i\in I}$ et pour tous $A_{i_1}, A_{i_2}, \cdots, A_{i_n} \in \mathfrak{B}_{\mathbb{R}}$ on a*

$$P\left(X_{i_1} \in A_{i_1}\,, \cdots, X_{i_n} \in A_{i_n}\right) = P(X_{i_1} \in A_{i_1}) \cdots P(X_{i_n} \in A_{i_n}) \tag{8.32}$$

*en notant*

$$P\left(X_{i_1} \in A_{i_1}\,, \cdots, X_{i_n} \in A_{i_n}\right) = P\left(\{X_{i_1} \in A_{i_1}\} \cap \cdots \cap \{X_{i_n} \in A_{i_n}\}\right).$$

**Exemples.** (i) Si $X, Y$ sont deux v.a. discrètes sur $(\Omega, \mathfrak{A}, P)$, à valeurs dans $\mathbb{N}$, elles sont indépendantes si et seulement

$$P(X = k\,,\, Y = \ell) = P(X = k)\; P(Y = \ell) \quad \forall\, k, \ell \in \mathbb{N}. \tag{8.33}$$

La condition (8.33) est évidemment nécessaire. Elle est suffisante puisqu'elle implique pour tous $A, B \subset \mathbb{N}$

$$\begin{aligned} P(X \in A\,,\, Y \in B) &= \sum_{(k,\ell)\in A\times B} P(X = k\,,\, Y = \ell) \\ &= \sum_{(k,\ell)\in A\times B} P(X = k)\; P(Y = \ell) \\ &= \left(\sum_{k\in A} P(X = k)\right)\left(\sum_{\ell\in B} P(Y = \ell)\right) \\ &= P(X \in A)\, P(Y \in B). \end{aligned}$$

(A noter que si $A, B$ sont infinis, on a utilisé le Th. 8.3 d'associativité des familles sommables de termes positifs).
(ii) Si $X_1, \cdots, X_{n+1}$ sont des v.a. à valeurs dans $\mathbb{N}$ indépendantes, $X_{n+1}$ et $X_1 + \cdots + X_n$ sont indépendantes. En effet, pour tous $k, \ell \in \mathbb{N}$, on a la partition

$$\{X_1 + \cdots + X_n = k\} \cap \{X_{n+1} = \ell\} = \bigcup_{\substack{0\leqslant k_1,\cdots,k_n\leqslant k \\ k_1+\cdots+k_n=k}} \{X_1 = k_1\} \cap \cdots \cap \{X_n = k_n\} \cap \{X_{n+1} = \ell\}$$

d'où

$$P(X_1 + \cdots + X_n = k\,,\, X_{n+1} = \ell) = \sum_{\substack{0\leqslant k_1,\cdots,k_n\leqslant k \\ k_1+\cdots+k_n=k}} P(X_1 = k_1) \ldots P(X_n = k_n)\, P(X_{n+1} = \ell)$$

qui s'écrit encore

$$P(X_1 + \cdots + X_n = k\,,\, X_{n+1} = \ell) = P(X_1 + \cdots + X_n = k)\; P(X_{n+1} = \ell).$$

**Définition 8. 15.** *Deux v.a. $X, Y$ sur sur $(\Omega, \mathfrak{A}, P)$ sont dites équidistribuées si elles ont la même loi :*

$$P_X = P_Y.$$

D'après le Lemme 8.7, $X, Y$ sont équidistribuées si et seulement si elles ont la même fonction de répartition :

$$F_X = F_Y.$$

Cela se généralise immédiatement à une famille quelconque de v.a.

### 8.4.2 Variables aléatoires discrètes.

Si $X$ est une v.a. discrète, les valeurs prises par $X$ forment une famille $(x_i)_{i\in I}$ avec $I$ dénombrable. On notera la loi de $X$

$$P_X = \sum_{i\in I} P(X = x_i)\,\delta_{x_i} \tag{8.34}$$

#### Variable aléatoire élémentaire de Bernouilli.

On appellera *v.a. élémentaire de Bernouilli* sur $(\Omega, \mathfrak{A}, P)$ une v.a. $X$ ne prenant que les valeurs 0 et 1 et dont la loi est $P_X = p\delta_1 + q\delta_0$ $(0 \leqslant p \leqslant 1)$.

On appelle *expérience élémentaire de Bernouilli* une expérience aléatoire modélisable par $(\{0,1\}, \mathfrak{P}(\{0,1\}), P)$ avec la mesure de probabilité $P$ sur $\{0,1\}$ définie par $P(\{1\}) = p$ et $P(\{0\}) = q = 1 - p$ $(0 \leqslant p \leqslant 1)$. On a $P = p\delta_1 + q\delta_0$. $P$ peut aussi être considérée comme une loi de probabilité sur $\mathbb{R}$.

Etant donnée une expérience élémentaire de Bernouilli, l'application $X : \{0,1\} \to \mathbb{R}$ définie par $X(0) = 0$ et $X(1) = 1$ est une v.a. élémentaire de Bernouilli sur $(\{0,1\}, \mathfrak{P}(\{0,1\}), P)$. Réciproquement, étant donnée une v.a. élémentaire de Bernouilli $X$ sur un espace probabilisé $(\Omega, \mathfrak{A}, P)$, l'expérience consistant à choisir au hasard un élément $\omega \in \Omega$ et à évaluer $X(\omega)$ est une expérience élémentaire de Bernouilli avec $p = P(X = 1)$ et $q = P(X = 0)$.

#### Processus de Bernouilli fini d'ordre $n$.

**Définition 8. 16.** *On appelle processus de Bernouilli fini d'ordre $n$ sur un espace probabilisé $(\Omega, \mathfrak{A}, P)$ une suite $(X_1, \cdots, X_n)$ de $n$ v.a. élémentaires de Bernouilli indépendantes et équidistribuées sur $(\Omega, \mathfrak{A}, P)$.*

**Exemple 1.** Soit l'expérience consistant en une suite finie de $n$ expériences élémentaires de Bernouilli identiques et indépendantes, modélisées chacune par la mesure de probabilité $P_0$ sur $\{0,1\}$ définie par $P_0(\{1\}) = p$ et $P_0(\{0\}) = q = 1 - p$, *i.e.* $P_0 = p\delta_1 + q\delta_0$ $(0 \leqslant p \leqslant 1)$. L'ensemble fondamental de cette expérience est

$$\Omega = \{\omega = (x_1, \ldots, x_n);\; x_i \in \{0,1\}\; \forall\, i\} = \{0,1\}^n.$$

Nous allons montrer qu'il existe sur $\mathfrak{P}(\Omega)$ une unique mesure de probabilité $P$ telle que

$$P(A_1 \times A_2 \times \cdots \times A_n) = P_0(A_1) \cdots P_0(A_n) \tag{8.35}$$

pour tous

$$A_1\,,\,A_2,\cdots,A_n\ \in \mathfrak{P}(\{0,1\})\ .$$

D'abord, si une telle mesure $P$ existe, comme pour tout $\omega=(x_1,\ldots,x_n)\in\Omega$ on a

$$\{\omega\}=\{x_1\}\times\ldots\times\{x_n\},$$

la condition (8.35) implique

$$P(\{\omega\})=P_0(\{x_1\})\ldots P_0(\{x_n\})=p^kq^{n-k} \tag{8.36}$$

où $k$ est le nombre de 1 dans $\omega$ et $n-k$ le nombre de 0.

Définissons donc $P(\{\omega\})$ pour tout $\omega=(x_1,\ldots,x_n)\in\Omega$ par la formule (8.36), et $P(A)$ pour tout $A\in\mathfrak{P}(\Omega)$ par

$$P(A)=\sum_{\omega\in A}P(\omega)\ . \tag{8.37}$$

Alors on obtient une mesure de probabilité $P$ sur $\mathfrak{P}(\Omega)$. En effet, d'abord la $\sigma$-additivité est trivialement réalisée d'après (8.37). Vérifions ensuite que

$$P(\Omega)=1.$$

Tout élément $\omega=(x_1,\ldots,x_n)\in\Omega$ est déterminé de façon unique par l'ensemble $\mathcal{I}_\omega\subset\mathfrak{P}(\{1,\cdots,n\})$ des indices $i$ tels que $x_i=1$, puisque qu'alors $x_i=0$ pour $i\notin\mathcal{I}_\omega$. Pour tout $\mathcal{I}\subset\mathfrak{P}(\{1,\cdots,n\})$, il existe un unique $\omega\in\Omega$ tel que $\mathcal{I}_\omega=\mathcal{I}$. Ainsi l'application

$$G\,:\,\Omega\to\mathfrak{P}(\{1,\cdots,n\}) \tag{8.38}$$

définie par $\omega\mapsto\mathcal{I}_\omega$ est une bijection. Par définition de $G$, dire que $\omega\in\Omega$ comporte $k$ fois le terme 1 ($0\leqslant k\leqslant n$) signifie que $|G(\omega)|=k$. Le nombre des tels $\omega$ est donc $C_n^k$. Or d'après (8.36), ces $\omega$ sont équiprobables, de probabilité $p^kq^{n-k}$ chacun. On a alors

$$P(\Omega)=\sum_{\omega\in\Omega}P(\omega)=\sum_{k=0}^{n}C_n^k\,p^kq^{n-k}=(p+q)^n=1.$$

Enfin la mesure de probabilité $P$ vérifie (8.35) puisque

$$\begin{aligned}
P(A_1\times A_2\times\cdots\times A_n) &= \sum_{(x_1,\cdots,x_n)\in A_1\times\cdots\times A_n} P(\{(x_1,\cdots,x_n)\})\\
&= \sum_{(x_1,\cdots,x_n)\in A_1\times\cdots\times A_n} P_0(\{x_1\})\cdots P_0(\{x_n\})\\
&= \left(\sum_{x_1\in A_1}P_0(\{x_1\})\right)\cdots\left(\sum_{x_n\in A_n}P_0(\{x_n\})\right)\\
&= P_0(A_1)\cdots P_0(A_n)
\end{aligned}$$

pour tous

$$A_1\,,\,A_2,\cdots,A_n\ \in\mathfrak{P}(\{0,1\})\ .$$

On a donc bien montré qu'il existe sur $\mathfrak{P}(\Omega)$ une unique mesure de probabilité $P$ vérifiant la condition (8.35).

Considérons maintenant pour tout $i = 1, \ldots, n$ l'application

$$\omega = (x_1, \ldots, x_n) \mapsto x_i$$

de $\Omega$ dans $\mathbb{R}$. Comme $P$ est définie sur $\mathfrak{P}(\Omega)$, cette application est une v.a. sur $(\Omega, \mathfrak{P}(\Omega), P)$. Notons la $X_i$. Si $A_i \in \mathfrak{P}(\{0,1\})$, on a

$$\begin{aligned} P(X_i \in A_i) &= P(\{0,1\} \times \{0,1\} \times \cdots \times A_i \times \cdots \times \{0,1\}) \\ &= P_0(A_i) \quad \text{d'après (8.35).} \end{aligned}$$

On en déduit immédiatement que la loi de $X_i$ est $P_{X_i} = p\delta_1 + q\delta_0 = P_0$, et $X_i$ est une v.a. élémentaire de Bernouilli. D'après (8.35), la suite $(X_1, \cdots, X_n)$ est une suite de v.a. indépendantes sur l'espace probabilisé $(\Omega, \mathfrak{P}(\Omega), P)$. Comme $X_1, \cdots, X_n$ ont la même loi, $(X_1, \cdots, X_n)$ est donc un processus de Bernouilli fini d'ordre $n$.

**Loi binomiale $\mathcal{B}(n, p)$.**

Soit $(X_1, \cdots, X_n)$ un processus de Bernouilli fini d'ordre $n$ sur un espace probabilisé $(\Omega, \mathfrak{A}, P)$. Soit $X$ la v.a.

$$X = X_1 + \ldots + X_n.$$

Les valeurs prises par $X$ sont $0, 1, \ldots, n$. Cette v.a. représente le *nombre de succès* dans la suite des expériences élémentaires associées aux v.a. $X_1, \cdots, X_n$. Calculons $P(X = k)$ pour $0 \leqslant k \leqslant n$. On a

$$\{X = k\} = \bigcup_{\substack{0 \leqslant k_1, \cdots, k_n \leqslant 1 \\ k_1 + \cdots + k_n = k}} \{X_1 = k_1\} \cap \cdots \cap \{X_n = k_n\}.$$

C'est une réunion disjointe, donc

$$P(X = k) = \sum_{\substack{0 \leqslant k_1, \cdots, k_n \leqslant 1 \\ k_1 + \cdots + k_n = k}} P(\{X_1 = k_1\} \cap \cdots \cap \{X_n = k_n\}).$$

Comme les v.a. $X_1, \cdots, X_n$ sont indépendantes,

$$P(X = k) = \sum_{\substack{0 \leqslant k_1, \cdots, k_n \leqslant 1 \\ k_1 + \cdots + k_n = k}} P(X_1 = k_1) \cdot \cdots \cdot P(X_n = k_n).$$

Or pour $1 \leqslant i \leqslant n$,

$$P(X_i = k_i) = \begin{cases} p & \text{si } k_i = 1 \\ q & \text{si } k_i = 0\,. \end{cases}$$

Donc

$$P(X = k) = \sum_{\substack{0 \leqslant k_1, \cdots, k_n \leqslant 1 \\ k_1 + \cdots + k_n = k}} p^k q^{n-k} = C_n^k\, p^k q^{n-k}.$$

La loi de $X$ est donc telle que

$$P(X = k) = C_n^k\, p^k q^{n-k}. \tag{8.39}$$

C'est la loi binomiale $\mathcal{B}(n,p)$ avec la définition suivante.

**Définition 8. 17.** *Soit $p$ , $0 \leqslant p \leqslant 1$ et $q = 1 - p$. On appelle loi binomiale de paramètres $n$ et $p$, et on note $\mathcal{B}(n,p)$ la loi de probabilité définie sur l'ensemble $\{0, \cdots, n\}$ par*

$$P(\{k\}) = C_n^k p^k q^{n-k} \quad \forall k \in \{0, \cdots, n\}. \tag{8.40}$$

*On considérera aussi cette loi comme une loi de probabilité sur $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$. Cela s'écrit aussi*

$$\mathcal{B}(n,p) = \sum_{k=0}^{n} C_n^k p^k q^{n-k} \delta_k.$$

Dans le cas de l'Exemple 1, par définition de l'application $G$ de (8.38), on a pour $0 \leqslant k \leqslant n$ :

$$X(\omega) = k \Leftrightarrow |G(\omega)| = k.$$

Le nombre des $\omega \in \Omega$ tels que $X(\omega) = k$ est donc $C_n^k$. Comme on l'a vu, ces $\omega$ sont équiprobables, de probabilité $p^k q^{n-k}$ chacun. Donc $P(X = k) = C_n^k p^k q^{n-k}$. On retrouve bien que $X$ suit la loi binomiale $\mathcal{B}(n,p)$.

**Exemple 2 : tirage avec remise.** Soit une urne contenant $N$ boules dont $N_1$ sont blanches et $N_2 = N - N_1$ ne sont pas blanches. On effectue $n$ tirages avec remise indépendants et on compte le nombre de fois où l'on a obtenu une blanche. On est dans le cas de l'exemple 1 : l'épreuve élémentaire de Bernouilli est ici le tirage d'une boule, avec la probabilité $p = \frac{N_1}{N}$ d'avoir une blanche et $q = 1 - p = \frac{N_2}{N}$ . Le nombre de fois où l'on aura obtenu une blanche est donc donné par la loi binomiale $\mathcal{B}(n,p)$, soit

$$P(X = k) = C_n^k p^k q^{n-k} = \frac{1}{N^n} C_n^k N_1^k N_2^{n-k}. \tag{8.41}$$

On peut modéliser différemment cette situation en introduisant un espace équiprobable. Soit $\Omega' = \mathcal{B}^n$ où $\mathcal{B}$ est l'ensemble de toutes les boules de l'urne. Un élément de $\Omega'$ est un $n$-uplet $(B_1, \ldots, B_n)$ où les $B_i$ sont des boules non nécessairement distinctes. Comme les tirages sont effectués avec remise, tous les $n$-uplets sont équiprobables, *i.e.* $\Omega'$ est équiprobable. Si $X$ est la v.a. sur $\Omega'$ "nombre de boules blanches du $n$-uplet", on a alors

$$P(X = k) = \frac{NCF}{NCP} = \frac{C_n^k N_1^k N_2^{n-k}}{N^n}$$

qui redonne (8.41).

**Loi hypergéométrique $\mathcal{H}(N, N_1, n)$.**

**Exemple : tirage sans remise.** On reprend l'exemple précédent, mais *sans remise.* On a donc une urne contenant $N$ boules dont $N_1$ sont blanches et $N_2 = N - N_1$ ne sont pas blanches. On effectue une suite de $n$ tirages sans remise ($n \leqslant N$), et soit $X$ le nombre de boules blanches obtenues. On désire définir $X$ comme v.a. sur un espace équiprobable. Soit $\Omega$ l'ensemble des parties à $n$ éléments de l'ensemble des $N$ boules. L'espace $\Omega$ est équiprobable, et $X$ est une v.a. sur $\Omega$.

Le nombre de parties à $n$ éléments de l'ensemble des $N$ boules comportant $k$ boules blanches est $C_{N_1}^{k} C_{N_2}^{n-k}$. On a donc :

$$P(X = k) = \frac{NCF}{NCP} = \frac{C_{N_1}^{k} C_{N_2}^{n-k}}{C_N^n}.$$

Notons que l'on doit avoir

$$0 \leqslant k \leqslant N_1, \quad 0 \leqslant n - k \leqslant N_2 = N - N_1$$

*i.e.*

$$k \leqslant \inf(N_1, n), \quad k \geqslant \sup(0, n - N_2).$$

$X$ suit donc la loi loi hypergéométrique $\mathcal{H}(N, N_1, n)$ de paramètres $N, N_1, n$ avec la définition suivante :

**Définition 8. 18.** *Soient $N, N_1, n$ des entiers tels que $0 \leqslant N_1 \leqslant N$ et $0 \leqslant n \leqslant N$. Soit $N_2 = N - N_1$. On appelle loi hypergéométrique $\mathcal{H}(N, N_1, n)$ de paramètres $N, N_1, n$ la loi de probabilité définie sur l'ensemble*

$$\Omega_{(N,N_1,n)} = \{k \in \mathbb{N};\ \sup(0, n - N_2) \leqslant k \leqslant \inf(N_1, n)\} \tag{8.42}$$

*par*

$$P(\{k\}) = \frac{C_{N_1}^{k} C_{N_2}^{n-k}}{C_N^n} \quad \forall k \in \Omega_{(N,N_1,n)}\,. \tag{8.43}$$

On notera que si $n \leqslant N_1$ et $n \leqslant N_2$, on a $\Omega_{(N,N_1,n)} = \{0, \cdots, n\}$. La loi hypergéométrique s'écrit encore :

$$P = \sum_{k \in \Omega_{(N,N_1,n)}} \frac{C_{N_1}^{k} C_{N_2}^{n-k}}{C_N^n}\, \delta_k.$$

La formule (8.43) définit bien une loi de probabilité sur $\Omega_{(N,N_1,n)}$ d'après le lemme suivant :

**Lemme 8. 8.**

$$\sum_{k \in \Omega_{(N,N_1,n)}} C_{N_1}^{k} C_{N_2}^{n-k} = C_N^n. \tag{8.44}$$

**Démonstration.**

Il suffit d'écrire le terme en $z^n$ dans $(1+z)^N = (1+z)^{N_1}(1+z)^{N_2}$. Plus précisément,

$$\begin{aligned}
(1+z)^{N_1}(1+z)^{N_2} &= \left(\sum_{0 \leqslant h_1 \leqslant N_1} C_{N_1}^{h_1} z^{h_1}\right)\left(\sum_{0 \leqslant h_2 \leqslant N_2} C_{N_2}^{h_2} z^{h_2}\right) \\
&= \sum_{\substack{0 \leqslant h_1 \leqslant N_1 \\ 0 \leqslant h_2 \leqslant N_2}} C_{N_1}^{h_1} C_{N_2}^{h_2} z^{h_1+h_2} \\
&= \sum_{j=0}^{N} \left(\sum_{\substack{0 \leqslant h_1 \leqslant N_1 \\ 0 \leqslant h_2 \leqslant N_2 \\ h_1+h_2=j}} C_{N_1}^{h_1} C_{N_2}^{h_2}\right) z^j\,.
\end{aligned}$$

Or

$$(1+z)^N = \sum_{j=0}^{N} C_N^j z^j \,,$$

Donc en comparant les termes en $z^n$ dans $(1+z)^N$ et $(1+z)^{N_1}(1+z)^{N_2}$, on obtient

$$C_N^n = \sum_{\substack{0 \leqslant h_1 \leqslant N_1 \\ 0 \leqslant h_2 \leqslant N_2 \\ h_1+h_2=n}} C_{N_1}^{h_1} C_{N_2}^{h_2} \,. \tag{8.45}$$

Cette formule n'est autre que (8.44) en écrivant $k = h_1$ et $h_2 = n - k$. □

**Processus de Bernouilli infini.**

**Définition 8. 19.** *On appelle processus de Bernouilli infini sur un espace probabilisé* $(\Omega, \mathfrak{A}, P)$ *une suite infinie* $(X_i)_{i\in\mathbb{N}^*}$ *de v.a. élémentaires de Bernouilli indépendantes et équidistribuées sur* $(\Omega, \mathfrak{A}, P)$.

**Exemple.** Considérons l'expérience consistant en une suite infinie d'expériences élémentaires de Bernouilli identiques et indépendantes, modélisées par la mesure de probabilité $P_0$ sur $\{0,1\}$ définie par $P_0(\{1\}) = p$ et $P_0(\{0\}) = q = 1-p$. L'ensemble fondamental est

$$\Omega = \{\omega = (x_n)_{n\in\mathbb{N}^*};\ x_n \in \{0,1\}\ \forall n \in \mathbb{N}^*\} = \{0,1\}^{\mathbb{N}^*}.$$

Soit $\mathfrak{R}$ l'ensemble des parties $X$ de $\Omega$ de la forme

$$X = A_1 \times A_2 \times \cdots \times A_{n(X)} \times \{0,1\} \times \{0,1\} \times \cdots$$

avec $A_1, \cdots, A_{n(X)} \in \mathfrak{P}(\{0,1\})$. Notons $\mathfrak{A}$ la plus petite $\sigma$-algèbre (pour l'inclusion dans $\mathfrak{P}(\Omega)$) contenant $\mathfrak{R}$. On admet que $\mathfrak{A} \neq \mathfrak{P}(\Omega)$ et qu'il existe sur $\mathfrak{A}$ une unique mesure de probabilité $P$ telle que

$$P(A_1 \times A_2 \times \cdots \times A_n \times \{0,1\} \times \{0,1\} \times \cdots) = P_0(A_1)\cdots P_0(A_n) \tag{8.46}$$

pour tout

$$A_1 \times A_2 \times \cdots \times A_n \times \{0,1\} \times \{0,1\} \times \cdots \in \mathfrak{R}$$

(voir [9] p.126, ex. 1 à 5). Pour tout $i \in \mathbb{N}^*$ soit $X_i$ la projection

$$\omega = (x_n)_{n\in\mathbb{N}^*} \mapsto x_i$$

de $\Omega$ dans $\mathbb{R}$. Par définition de $\mathfrak{A}$, $X_i$ est une v.a. et, d'après (8.46), la suite $(X_i)_{i\in\mathbb{N}^*}$ est une suite de v.a. indépendantes sur l'espace probabilisé $(\Omega, \mathfrak{A}, P)$. Comme les $X_i$ ont la même loi, $(X_i)_{i\in\mathbb{N}^*}$ est donc un processus de Bernouilli infini.

**Loi binomiale négative $\mathcal{B}^-(r,p)$.**

**Définition 8. 20.** *Soit* $p$ , $0 < p < 1$ *et* $r \in \mathbb{N}^*$. *On appelle loi binomiale négative de paramètres* $r$ *et* $p$, *et on note* $\mathcal{B}^-(r,p)$, *la loi de probabilité définie sur l'ensemble* $\Omega_r = r + \mathbb{N}$ *par*

$$P(\{r+k\}) = \binom{-r}{k} p^r\,(-q)^k \quad \forall k \in \mathbb{N} \tag{8.47}$$

*avec* $q = 1 - p$, $\binom{-r}{0} = 1$, *et pour* $k \geqslant 1$ :

$$\begin{aligned}\binom{-r}{k} &= \frac{(-r)(-r-1)\cdots(-r-(k-1))}{k!} \\ &= (-1)^k \frac{r(r+1)\cdots(r+k-1)}{k!} \\ &= (-1)^k C_{r+k-1}^k.\end{aligned}$$

On notera que la formule

$$\binom{-r}{k} = (-1)^k C_{r+k-1}^k$$

est encore valable pour $k = 0$.

La formule (8.47) définit bien une loi de probabilité sur $r + \mathbb{N}$. En effet, on sait que pour $|x| < 1$ on a le développement en série entière

$$(1+x)^{-r} = \sum_{k=0}^{+\infty} \binom{-r}{k} x^k,$$

donc comme $0 < p, q < 1$,

$$\sum_{k=0}^{+\infty} \binom{-r}{k} p^r (-q)^k = p^r \sum_{k=0}^{+\infty} \binom{-r}{k} (-q)^k = p^r (1-q)^{-r} = p^r p^{-r} = 1\,. \qquad (8.48)$$

La loi binomiale négative $\mathcal{B}^-(r,p)$ s'écrit :

$$P = \sum_{k=0}^{+\infty} \binom{-r}{k} p^r (-q)^k \delta_{r+k}.$$

**Exemple.** Soit $(X_i)_{i\in\mathbb{N}^*}$ un processus de Bernouilli infini sur un espace probabilisé $(\Omega, \mathfrak{A}, P)$. On suppose $0 < p < 1$. Soit $r \in \mathbb{N}^*$ et $T_r$ l'application de $\Omega$ dans $\mathbb{R}\cup\{+\infty\}$ définie pour tout $\omega \in \Omega$ par

$$T_r(\omega) = \inf\{n \in \mathbb{N}^* \,;\, X_1(\omega) + \cdots + X_n(\omega) = r\}$$

si $\{n \in \mathbb{N}^* \,;\, X_1(\omega) + \cdots + X_n(\omega) = r\} \neq \varnothing$ et $T_r(\omega) = +\infty$ sinon. $T_r$ est à valeurs dans $(r + \mathbb{N}) \cup \{+\infty\}$. On a pour $k \in \mathbb{N}$

$$(T_r)^{-1}(\{r+k\}) = \{T_r = r+k\} = \{X_{r+k} = 1\} \cap \{X_1 + \cdots + X_{r+k-1} = r-1\} \qquad (8.49)$$

donc $(T_r)^{-1}(\{r+k\}) \in \mathfrak{A}$ puisque $X_{r+k}$ et $X_1 + \cdots + X_{r+k-1}$ sont des v.a. On a également

$$(T_r)^{-1}(\{+\infty\}) = \left(\bigcup_{k\in\mathbb{N}} (T_r)^{-1}(\{r+k\})\right)^c \in \mathfrak{A}.$$

Maintenant, l'équation (8.49) donne puisque $X_{r+k}$ et $X_1 + \cdots + X_{r+k-1}$ sont des v.a. indépendantes et que $X_1 + \cdots + X_{r+k-1}$ suit la loi binomiale $\mathcal{B}(r+k-1, p)$ :

$$\begin{aligned} P(T_r = r+k) &= P(X_{r+k} = 1)\, C_{r+k-1}^{r-1} p^{r-1} q^k \\ &= C_{r+k-1}^{r-1} p^r q^k \\ &= C_{r+k-1}^{k} p^r q^k \\ &= \binom{-r}{k} p^r (-q)^k , \end{aligned}$$

avec les notations de la définition 8.20. En particulier $\sum_{k=0}^{+\infty} P(T_r = r+k) = 1$ d'après (8.48) et donc $P(T_r = +\infty) = 1 - \sum_{k=0}^{+\infty} P(T_r = r+k) = 0$. On peut ainsi écarter la valeur $+\infty$ et considérer $T_r$ comme une v.a. à valeurs dans $r + \mathbb{N}$. La loi de $T_r$ est alors la loi binomiale négative $\mathcal{B}^-(r, p)$.

**Loi géométrique $\mathcal{G}(p)$.**

Pour $r = 1$, la loi binomiale négative $\mathcal{B}^-(1, p)$ donne $P(\{1+k\}) = pq^k$ si $k \in \mathbb{N}$. C'est la loi géométrique avec la définition suivante :

**Définition 8. 21.** *Soit $p$ , $0 < p < 1$. On appelle loi géométrique de paramètre $p$ et on note $\mathcal{G}(p)$ la loi de probabilité définie sur l'ensemble $\Omega = \mathbb{N}^*$ par*

$$P(\{k\}) = pq^{k-1} \quad \forall k \in \mathbb{N}^* \tag{8.50}$$

*avec $q = 1 - p$.*

La formule (8.50) définit bien une loi de probabilité sur $\mathbb{N}^*$ puisque que c'est la loi de $T_r$ pour $r = 1$. Mais on le voit aussi directement :

$$\sum_{k=1}^{+\infty} pq^{k-1} = \frac{p}{1-q} = 1.$$

La loi $\mathcal{G}(p)$ s'écrit :

$$P = \sum_{k=1}^{+\infty} pq^{k-1} \delta_k.$$

C'est la loi des prédateurs car un prédateur, par exemple un lion, s'arrête de chasser et dévore sa proie au premier succès.

### 8.4.3 Variables aléatoires continues.

Soit $X$ une v.a. continue sur un espace probabilisé $(\Omega, \mathfrak{A}, P)$ et $P_X$ sa loi de probabilité. Si $P_X$ est définie par une densité de probabilité $g$, on dit que $X$ est une v.a. de densité $g$ sur $\mathbb{R}$. Une v.a. a pour densité $g$ si et seulement si

$$P(a \leqslant X \leqslant b) = \int_a^b g(t)\, dt \quad \forall a, b \in \mathbb{R}\ a < b.$$

**Théorème 8. 2.** *Soit $X : \Omega \to \mathbb{R}$ une variable aléatoire de densité $g$, et soient $\alpha, \beta \in \mathbb{R}$ avec $\alpha \neq 0$. Alors la variable aléatoire $Y = \alpha X + \beta$ a pour densité la fonction $h$ définie par*

$$h(s) = \frac{1}{|\alpha|} g\left(\frac{s-\beta}{\alpha}\right) \quad \forall s \in \mathbb{R}. \tag{8.51}$$

**Démonstration.**

Il faut déterminer une fonction $h$ telle que

$$P(u \leqslant Y \leqslant v) = \int_u^v h(s)\, ds \quad \forall u, v \in \mathbb{R},\ u < v.$$

Or $u \leqslant Y \leqslant v$ s'écrit

$$\begin{cases} \frac{u-\beta}{\alpha} \leqslant X \leqslant \frac{v-\beta}{\alpha} & \text{si } \alpha > 0 \\ \frac{u-\beta}{\alpha} \geqslant X \geqslant \frac{v-\beta}{\alpha} & \text{si } \alpha < 0 \end{cases}$$

d'où

$$\begin{aligned} P(u \leqslant Y \leqslant v) &= \varepsilon \int_{\frac{u-\beta}{\alpha}}^{\frac{v-\beta}{\alpha}} g(t)\, dt \\ &= \varepsilon \int_u^v g(\frac{s-\beta}{\alpha}) \frac{ds}{\alpha} \\ &= \frac{1}{|\alpha|} \int_u^v g(\frac{s-\beta}{\alpha})\, ds \end{aligned}$$

où $\varepsilon = \pm 1$ suivant le signe de $\alpha$ et l'on a effectué le changement de variable

$$t = \frac{s-\beta}{\alpha}.$$

On en déduit que la densité de $Y$ est la fonction définie par

$$h(s) = \frac{1}{|\alpha|} g(\frac{s-\beta}{\alpha}) \ \forall s \in \mathbb{R}.$$

□

**Application à la loi normale.**

Si $X$ est une v.a. qui suit la loi normale centrée réduite $\mathcal{N}(0,1)$, de densité $g(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$, la v.a. $Y = \sigma X + m$ ($\sigma > 0$, $m \in \mathbb{R}$) a pour densité

$$h(s) = \frac{1}{\sigma} g(\frac{s-m}{\sigma}) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(s-m)^2}{2\sigma^2}}.$$

C'est la densité de la loi loi normale générale $\mathcal{N}(m,\sigma)$, donc $Y$ suit $\mathcal{N}(m,\sigma)$. Réciproquement, si $Y$ est une v.a. qui suit $\mathcal{N}(m,\sigma)$, la v.a. $X = \frac{Y-m}{\sigma}$ suit la loi normale centrée réduite $\mathcal{N}(0,1)$.

## 8.5 Moments d'une variable aléatoire.

### 8.5.1 Rappels sur les familles sommables.

Nous rappelons ici sans démonstration quelques résultats fondamentaux sur les familles sommables de réels. Les démonstrations figurent dans l'Appendice 2 de ce chapitre.

**Familles de réels.**

Soit $I$ un ensemble. Une *famille* de réels est une application $i \mapsto u_i$ de $I$ dans $\mathbb{R}$. On la note $(u_i)_{i\in I}$. Si $I = \mathbb{N}$, la famille est simplement une suite $(u_n)_{n\in\mathbb{N}}$. Si $I = \varnothing$, on convient de poser $\sum_{i\in I} u_i = 0$.

**Somme d'une famille de réels $\geqslant 0$.**

Soit $(u_i)_{i\in I}$ une famille d'éléments de $[0, +\infty]$. On pose :

$$\sum_{i\in I} u_i = \sup_{\substack{F\subset I \\ F \text{fini}}} \sum_{i\in F} u_i \quad \in \quad [0, +\infty]\,. \tag{8.52}$$

**Théorème 8. 3.** *Soit $(u_i)_{i\in I}$ une famille d'éléments de $[0, +\infty]$. Si $I = \bigcup_{\lambda\in\Lambda} I_\lambda$ est une partition quelconque de $I$, on a alors :*

$$\sum_{i\in I} u_i = \sum_{\lambda\in\Lambda}\left(\sum_{i\in I_\lambda} u_i\right) \tag{8.53}$$

En particulier, dans le cas d'une famille *produit*, encore appelée famille *double* $(u_{i,j})_{(i,j)\in I\times J}$ de réels $\geqslant 0$, pour vérifier que $\sum_{(i,j)\in I\times J} u_{i,j} < +\infty$, il suffit de vérifier que *l'une* des deux sommes

$$\sum_{i\in I}\left(\sum_{j\in J} u_{i,j}\right) \text{ ou } \sum_{j\in J}\left(\sum_{i\in I} u_{i,j}\right)$$

est finie.

**Famille sommable de réels.**

**Définition 8. 22.** *Une famille $(u_i)_{i\in I}$ de réels est dite sommable de somme $S$ si pour tout $\varepsilon > 0$ il existe une partie finie $F_0$ de $I$ ayant la propriété suivante : pour toute partie finie $F$ de $I$ contenant $F_0$, on a*

$$\left|S - \sum_{i\in F} u_i\right| < \varepsilon. \tag{8.54}$$

**Proposition 8. 2.** *Une famille $(u_i)_{i\in I}$ de réels $\geqslant 0$ est sommable si et seulement si la somme (8.52) est $< +\infty$ et sa somme $S$ est alors la somme (8.52).*

Pour cette raison, on note aussi $\sum_{i\in I} u_i$ la somme $S$ d'une famille sommable quelconque.

**Espace vectoriel des familles sommables de réels.**

Fixons l'ensemble d'indices $I$. Considérons l'ensemble noté $\mathbb{R}^I$ de toutes les familles de réels ayant l'ensemble d'indices $I$. Soient $(u_i)_{i\in I}$, $(v_i)_{i\in I} \in \mathbb{R}^I$. On définit la famille somme de ces deux familles comme étant la famille $(u_i + v_i)_{i\in I} \in \mathbb{R}^I$. On définit aussi la famille produit par $\lambda \in \mathbb{R}$ de la famille $(u_i)_{i\in I} \in \mathbb{R}^I$ comme étant la famille $(\lambda u_i)_{i\in I}$. Il est facile de vérifier que muni de ces deux opérations, $\mathbb{R}^I$ est un espace vectoriel réel.

Supposons maintenant les deux familles $(u_i)_{i\in I}$, $(v_i)_{i\in I} \in \mathbb{R}^I$ sommables, de sommes respectives $S$ et $T$. Soit $\varepsilon > 0$ quelconque. Il existe une partie finie $F_0$ de $I$ ayant la propriété suivante : pour toute partie finie $F$ de $I$ contenant $F_0$, on a

$$\left|S - \sum_{i\in F} u_i\right| < \frac{\varepsilon}{2}.$$

De même, il existe une partie finie $G_0$ de $I$ ayant la propriété suivante : pour toute partie finie $F$ de $I$ contenant $G_0$, on a

$$\left|T - \sum_{i\in F} u_i\right| < \frac{\varepsilon}{2}.$$

Soit $H_0 = F_0 \cup G_0$. C'est une partie finie de $I$, et pour toute partie finie $F$ de $I$ contenant $H_0$, on a

$$\left|(S+T) - \sum_{i\in F} (u_i + v_i)\right| < \varepsilon.$$

Donc la famille $(u_i + v_i)_{i\in I}$ est sommable de somme $S + T$.

Il est par ailleurs immédiat que la famille $(\lambda u_i)_{i\in I}$ est sommable de somme $\lambda S$ pour tout $\lambda \in \mathbb{R}$. L'ensemble des familles sommables est donc un sous-espace vectoriel de l'espace vectoriel $\mathbb{R}^I$, et l'application qui à toute famille sommable associe sa somme est une forme linéaire sur ce sous-espace.

**Famille sommable et famille absolument sommable de réels.**

**Définition 8. 23.** *Une famille $(u_i)_{i\in I}$ de réels est dite absolument sommable si la famille $(|u_i|)_{i\in I}$ est sommable, i.e.*

$$\sum_{i\in I} |u_i| < +\infty. \tag{8.55}$$

Une propriété fondamentale est la suivante :

**Théorème 8. 4.** *Une famille de réels est sommable si et seulement si elle est absolument sommable, i.e. vérifie (8.55). On a alors*

$$\left|\sum_{i\in I} u_i\right| \leqslant \sum_{i\in I} |u_i|. \tag{8.56}$$

En particulier, dans le cas d'une suite $(u_n)_{n\in\mathbb{N}}$, la famille $(u_n)_{n\in\mathbb{N}}$ est sommable si et seulement si la série $\sum_{n=0}^{\infty} u_n$ est absolument convergente. La somme $S$ de la famille sommable est alors la somme de la série. En effet, pour tout $\varepsilon > 0$ il existe une partie finie $F_0 \subset \mathbb{N}$ telle que pour toute partie finie $F \subset \mathbb{N}$ contenant $F_0$ on ait $\left|S - \sum_{n\in F} u_n\right| < \varepsilon$. Alors $\left|S - \sum_{n=0}^{N} u_n\right| < \varepsilon \ \forall N \geqslant \sup F_0$ puisque $F = \{1, \cdots, N\} \supset F_0$. Comme $\varepsilon > 0$ est arbitraire, $S$ est la somme de la série. La notion de famille sommable généralise donc celle de série absolument convergente.

Notons aussi que si $(u_i)_{i\in I}$ est une famille sommable, toute sous-famille $(u_i)_{i\in J}$ avec $J \subset I$ est sommable. C'est une conséquence du Théorème 8.4 puisque $\sum_{i\in J} |u_i| \leqslant \sum_{i\in I} |u_i| < +\infty$.

**Théorème d'associativité pour les familles sommables de réels.**

Soit $(u_i)_{i\in I}$ une famille de réels et $I = \bigcup_{\lambda\in\Lambda} I_\lambda$ une partition de $I$. D'après (8.55) et (8.53), pour que la famille $(u_i)_{i\in I}$ soit sommable, il faut et il suffit que

$$\sum_{i\in I} |u_i| = \sum_{\lambda\in\Lambda} \left(\sum_{i\in I_\lambda} |u_i|\right) < +\infty\,. \tag{8.57}$$

**Théorème 8. 5.** *Soit $(u_i)_{i\in I}$ une famille sommable de réels et $S = \sum_{i\in I} u_i$ sa somme. Soit $I = \bigcup_{\lambda\in\Lambda} I_\lambda$ une partition de $I$. Alors pour chaque $\lambda \in \Lambda$ la famille $(u_i)_{i\in I_\lambda}$ est sommable de somme $s_\lambda = \sum_{i\in I_\lambda} u_i$, et la famille $(s_\lambda)_{\lambda\in\Lambda}$ est sommable de somme $S$, i.e.*

$$\sum_{i\in I} u_i = \sum_{\lambda\in\Lambda} \left(\sum_{i\in I_\lambda} u_i\right) \tag{8.58}$$

## 8.5.2 Cas d'une loi discrète.

Dans toute cette section, on ne considère que des v.a. *discrètes* sur un espace probabilisé $(\Omega, \mathfrak{A}, P)$. Les valeurs prises par une telle v.a. $X$ forment une famille $(x_i)_{i\in I}$ avec $I$ dénombrable (fini ou infini). La loi de $X$ est notée comme en (8.34) $P_X = \sum_{i\in I} P(X = x_i)\,\delta_{x_i}$.

**Espérance.**

**Définition 8. 24.** *On dit que la v.a. $X$ est sommable si la famille*

$$(P(X = x_i)\, x_i)_{i\in I} \tag{8.59}$$

*est sommable. Si $X$ est sommable, on appelle espérance de $X$ et on note $E(X)$ la somme de la famille (8.59) :*

$$E(X) = \sum_{i\in I} P(X = x_i)\, x_i \tag{8.60}$$

**Proposition 8. 3.** *Si $X, Y$ sont des v.a. discrètes sommables sur $(\Omega, \mathfrak{A}, P)$, il en est de même des v.a. $\alpha X$ ($\alpha \in \mathbb{R}$) et $X + Y$ et on a*

$$E(\alpha X) = \alpha E(X) \quad \forall \alpha \in \mathbb{R}\,, \tag{8.61}$$
$$E(X + Y) = E(X) + E(Y)\,. \tag{8.62}$$

**Démonstration.**

Considérons d'abord $X+Y$. On a

$$P_X = \sum_{i\in I} P(X=x_i)\,\delta_{x_i}\,, \quad P_Y = \sum_{j\in J} P(Y=y_j)\,\delta_{y_j}\,.$$

Par définition,

$$E(X) = \sum_{i\in I} P(X=x_i)\,x_i, \quad E(Y) = \sum_{j\in J} P(Y=y_j)\,y_j,$$

les deux familles étant sommables. La v.a. $X+Y$ prend les valeurs $x_i+y_j$, et l'on a

$$P_{X+Y} = \sum_{(i,j)\in I\times J} P(X=x_i, Y=y_j)\,\delta_{x_i+y_j}\,.$$

Dans cette expression $P(X=x_i, Y=y_j)$ désigne $P(\{X=x_i\}\cap\{Y=y_j\})$ et les $x_i+y_j$ ne sont pas forcément deux-à-deux distincts. Il s'agit de montrer que la famille double

$$\left(P(X=x_i, Y=y_j)\,(x_i+y_j)\right)_{(i,j)\in I\times J}\,, \tag{8.63}$$

est sommable de somme $E(X)+E(Y)$. Considérons d'abord la famille double

$$\left(P(X=x_i, Y=y_j)\,x_i\right)_{(i,j)\in I\times J}\,. \tag{8.64}$$

Comme

$$\begin{aligned}\bigcup_{j\in J}(\{X=x_i\}\cap\{Y=y_j\}) &= \{X=x_i\}\cap\bigcup_{j\in J}\{Y=y_j\} \\ &= \{X=x_i\}\cap\Omega = \{X=x_i\}\end{aligned}$$

et que la réunion est disjointe, on a

$$\sum_j P(X=x_i, Y=y_j) = P(X=x_i) \quad \forall i\in I. \tag{8.65}$$

Donc

$$\sum_i\sum_j P(X=x_i, Y=y_j)\,|x_i| = \sum_i P(X=x_i)\,|x_i| < +\infty. \tag{8.66}$$

Cela implique que la famille double (8.64) est sommable. La famille étant sommable, sa somme est d'après le théorème d'associativité Th.8.5 et (8.65) :

$$\begin{aligned}\sum_{i,j} P(X=x_i, Y=y_j)\,x_i &= \sum_{i\in I}\sum_{j\in J} P(X=x_i, Y=y_j)\,x_i. \\ &= \sum_{i\in I} P(X=x_i)\,x_i \\ &= E(X).\end{aligned}$$

De même, de façon analogue à (8.65) on a

$$\sum_i P(X=x_i, Y=y_j) = P(Y=y_j) \quad \forall j\in J. \tag{8.67}$$

La famille double

$$(P(X = x_i, Y = y_j)\, y_j)_{(i,j)\in I\times J} \tag{8.68}$$

est sommable et sa somme est

$$\begin{aligned}\sum_{i,j} P(X = x_i, Y = y_j)\, y_j &= \sum_{j\in J}\sum_{i\in I} P(X = x_i, Y = y_j)\, y_j \\ &= \sum_{j\in J} P(Y = y_j)\, y_j \\ &= E(Y).\end{aligned}$$

Or la famille (8.63) est la somme des deux familles (8.64) et (8.68). Donc elle est sommable, et sa somme est

$$\sum_{i,j} P(X = x_i, Y = y_j)\, x_i + \sum_{i,j} P(X = x_i, Y = y_j)\, y_j = E(X) + E(Y).$$

Cela prouve que $X + Y$ est sommable et que $E(X + Y) = E(X) + E(Y)$.

Considérons maintenant $\alpha X$ . La loi de $\alpha X$ est

$$P_{\alpha X} = \sum_{i\in I} P(X = x_i)\, \delta_{\alpha x_i}\, .$$

La famille $(P(X = x_i)\, \alpha x_i)_{i\in I}$ est sommable de somme $\alpha \sum_{i\in I} P(X = x_i)\, x_i$ . Donc $\alpha X$ est sommable et $E(\alpha X) = \alpha\, E(X)$ . □

**Corollaire.** *Les v.a. sommables forment un sous-espace vectoriel de l'espace vectoriel des v.a. réelles contenant les constantes, et l'application $E : X \mapsto E(X)$ est une forme linéaire sur ce sous-espace.*

**Démonstration.**

Cela résulte immédiatement des définitions et de la Proposition 8.3. □

**Définition 8. 25.** *Soit $X$ une v.a. sommable. La v.a. $Y = X - E(X)$ est appelée la variable centrée.*

L'espérance étant une forme linéaire et l'espérance d'une v.a. constante étant égale à cette constante, on a

$$E(Y) = E(X) - E(E(X)) = E(X) - E(X) = 0.$$

L'espérance de la v.a. centrée est donc nulle.

**Proposition 8. 4.** *Soient $X, Y$ deux v.a. telles que $0 \leqslant X \leqslant Y$. Si $Y$ est sommable, il en est de même de $X$ et $E(X) \leqslant E(Y)$.*

**Démonstration.**

Les valeurs prises par $X$ sont la famille $(x_i)_{i\in I}$ et $x_i \geqslant 0$ pour tout $i$. Les valeurs prises par $Y$ sont la famille $(y_j)_{j\in J}$, et $y_j \geqslant 0$ pour tout $j$ puisque pour tout $j$ il existe $\omega \in \Omega$ tel que $Y(\omega) = y_j$ et alors $y_j = Y(\omega) \geqslant X(\omega) \geqslant 0$. Par hypothèse, la famille $(P(Y = y_j)\, y_j)_{j\in J}$ est sommable de somme $E(Y) = \sum_{j\in J} P(Y = y_j)\, y_j$. Donc d'après (8.67)

$$E(Y) = \sum_j \sum_i P(X = x_i, Y = y_j)\, y_j.$$

Pour tous $i, j$, si $P(X = x_i, Y = y_j) \neq 0$, il existe $\omega \in \Omega$ tel que $X(\omega) = x_i$ et $Y(\omega) = y_j$, et alors $x_i = X(\omega) \leqslant Y(\omega) = y_j$. On en déduit que pour tous $i, j$

$$P(X = x_i, Y = y_j)\, x_i \leqslant P(X = x_i, Y = y_j)\, y_j\,,$$

donc

$$\sum_{i,j} P(X = x_i, Y = y_j)\, x_i \leqslant \sum_{i,j} P(X = x_i, Y = y_j)\, y_j = E(Y) < +\infty\,.$$

Cela prouve que la famille double

$$(P(X = x_i, Y = y_j)\, x_i)_{(i,j)\in I\times J}$$

est sommable de somme $\leqslant E(Y)$. Or en utilisant (8.65) il vient

$$\sum_{i,j} P(X = x_i, Y = y_j)\, x_i = \sum_i P(X = x_i)\, x_i = E(X).$$

Donc $X$ est sommable et $E(X) \leqslant E(Y)$. □

**Ecart-type.**

**Définition 8. 26.** *Soit $X$ une v.a. On dit que $X$ est de carré sommable si la v.a. $X^2$ est sommable.*

Si la loi de $X$ est (8.34)

$$P_X = \sum_{i\in I} P(X = x_i)\, \delta_{x_i}\,,$$

la loi de $X^2$ est

$$P_{X^2} = \sum_{i\in I} P(X = x_i)\, \delta_{x_i^2}\,,$$

où dans cette dernière expression, les $x_i^2$ ne sont pas nécessairement distincts.

**Lemme 8. 9.** *(i) Le produit $XY$ de deux v.a. $X, Y$ de carré sommable est sommable.*
*(ii) Si $X, Y$ sont deux v.a. de carré sommable, il en est de même des v.a. $\alpha X$ ($\alpha \in \mathbb{R}$) et $X + Y$.*
*(iii) Une v.a. $X$ de carré sommable est sommable.*

**Démonstration.**
(i) On a $|XY| \leqslant \frac{1}{2}(X^2 + Y^2)$, donc $|XY|$, et par conséquent $XY$, est sommable d'après la Prop. 8.4.
(ii) D'après (i), $XY$ est sommable. Or $(X+Y)^2 = X^2 + Y^2 + 2XY$, donc $X + Y$ est de carré sommable. Le cas de $\alpha X$ est immédiat.
(iii) La v.a. $Y = 1$ est de carré sommable. Donc $X = XY$ est sommable. □

**Corollaire.** *L'ensemble des v.a. de carré sommable est un sous-espace vectoriel de l'espace vectoriel des v.a. sommables contenant les constantes.*

**Démonstration.**

D'après lemme, $X+Y$ et $\alpha X$ ($\alpha \in \mathbb{R}$) sont des v.a. de carré sommable. De plus toute v.a. constante est trivialement de carré sommable. □

**Définition 8. 27.** *Si $X$ est de carré sommable, $E(X^2)$ est appelé moment d'ordre 2 de $X$.*

Il est clair que le moment d'ordre 2 d'une v.a. de carré sommable est $\geqslant 0$.

**Lemme 8. 10.** *Soit $X$ une v.a. de carré sommable et $Y = X - E(X)$ la variable centrée. Alors $Y$ est de carré sommable et son moment d'ordre 2 est*

$$E(Y^2) = E(X^2) - (E(X))^2.$$

**Démonstration.**

$Y$ est de carré sommable d'après le Corollaire du Lemme 8.9. Or

$$Y^2 = (X - E(X))^2 = X^2 - 2E(X)\,X + (E(X))^2.$$

Donc (Prop 8.3)

$$\begin{aligned} E(Y^2) &= E(X^2) - 2E(E(X)\,X) + E((E(X))^2) \\ &= E(X^2) - 2(E(X))^2 + (E(X))^2 \\ &= E(X^2) - (E(X))^2. \end{aligned}$$

□

**Définition 8. 28.** *Soit $X$ une v.a. de carré sommable. On appelle variance de $X$ et on note $var(X)$ le moment d'ordre 2 de la variable centrée :*

$$var(X) = E(X^2) - (E(X))^2.$$

*On appelle écart-type de $X$ et on note $\sigma(X)$ la racine carrée de la variance de $X$ :*

$$\sigma(X) = \sqrt{var(X)} = \sqrt{E(X^2) - (E(X))^2}.$$

*Si $X, Y$ sont deux v.a. de carré sommable, on appelle covariance de $X$ et $Y$ et on note $cov(X,Y)$ la quantité $E(XY) - E(X)E(Y)$.*

**Proposition 8. 5.** *Soient $X, Y$ des v.a. de carré sommable.*
*(i) $var(\alpha X) = \alpha^2 var(X) \quad \forall \alpha \in \mathbb{R}$.*
*(ii) $var(X+Y) = var(X) + var(Y) + 2cov(X,Y)$.*
*(iii) Si les v.a. $X$ et $Y$ sont indépendantes on a*

$$var(X+Y) = var(X) + var(Y).$$

**Démonstration.**

(i) $\mathrm{var}(\alpha X) = E(\alpha^2 X^2) - (E(\alpha X))^2 = \alpha^2 E(X^2) - (\alpha E(X))^2 = \alpha^2 \mathrm{var}(X)$ d'après la prop. 8.3 .
(ii) On a $E((X+Y)^2) = E(X^2+Y^2+2XY)$. Comme $X$ et $Y$ sont de carré sommable, $XY$ est sommable et alors

$$E((X+Y)^2) = E(X^2) + E(Y^2) + E(2XY) = E(X^2) + E(Y^2) + 2E(XY).$$

D'autre part

$$(E(X+Y))^2 = (E(X)+E(Y))^2 = E(X)^2 + E(Y)^2 + 2E(X)E(Y).$$

Donc $\operatorname{var}(X+Y) = \operatorname{var}(X) + \operatorname{var}(Y) + 2\operatorname{cov}(X,Y)$.
(iii) Si $X$ et $Y$ sont indépendantes, on a

$$P(X = x_i, Y = y_j) = P(X = x_i)\, P(Y = y_j)$$

pour tous $i, j$ d'où

$$\begin{aligned} E(XY) &= \sum_{i,j} P(X = x_i, Y = y_j)\, x_i y_j \\ &= \sum_{i,j} P(X = x_i)\, P(Y = y_j)\, x_i y_j \\ &= \left(\sum_i P(X = x_i)\, x_i\right)\left(\sum_j P(Y = y_j)\, y_j\right) \\ &= E(X)\, E(Y) \end{aligned}$$

donc $\operatorname{cov}(X,Y) = 0$ et alors $\operatorname{var}(X+Y) = \operatorname{var}(X) + \operatorname{var}(Y)$. □
**Remarque.** On a pour $X$ v.a. sommable et $\alpha, \beta \in \mathbb{R}$ d'après la Prop. 8.3 :

$$E(\alpha X + \beta) = \alpha\, E(X) + \beta. \tag{8.69}$$

Si $X$ est de carré sommable, on a d'après la Prop. 8.5 en notant que $\operatorname{cov}(\alpha X, \beta) = 0$ et que la variance d'une constante est nulle :

$$\operatorname{var}(\alpha X + \beta) = \alpha^2 \operatorname{var}(X). \tag{8.70}$$

Pour $X, Y$ v.a. de carré sommable et $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ on a :

$$\operatorname{cov}(\alpha X + \beta, \gamma Y + \delta) = \alpha\gamma \operatorname{cov}(X,Y). \tag{8.71}$$

**Variable centrée réduite.**

**Définition 8. 29.** *Soit $X$ une v.a. de carré sommable. La v.a.*

$$Y = \frac{X - E(X)}{\sigma(X)}$$

*est appelée variable centrée réduite associée à $X$.*

La variable centrée réduite a pour espérance 0 et pour écart-type 1.

**Exemples.**

**Loi binomiale $\mathcal{B}(n,p)$.**

Soit $X$ une v.a. qui suit la loi $\mathcal{B}(n,p)$. Elle est évidemment de carré sommable puisqu'elle ne prend qu'un nombre fini de valeurs. On a $X = X_1 + \ldots + X_n$, où les $X_i$ sont des v.a. élémentaires de Bernouilli indépendantes. Or une v.a. élémentaire de

Bernouilli est de carré sommable, son espérance est $p$ et sa variance est $p - p^2 = pq$. Donc

$$\begin{aligned} E(X) &= E(X_1) + \ldots + E(X_n) = np & (8.72)\\ \operatorname{var}(X) &= \operatorname{var}(X_1) + \ldots + \operatorname{var}(X_n) = npq\,. & (8.73)\end{aligned}$$

On peut aussi calculer $E(X)$ directement :

$$\begin{aligned} E(X) &= \sum_{k=0}^{n} P(X=k)\,k \\ &= \sum_{k=0}^{n} k\, C_n^k\, p^k q^{n-k} \\ &= p \sum_{k=1}^{n} k\, C_n^k\, p^{k-1} q^{n-k} \\ &= p \left[\frac{d}{dt} \sum_{k=0}^{n} C_n^k\, t^k q^{n-k}\right]_{t=p} \\ &= p \left[\frac{d}{dt}(t+q)^n\right]_{t=p} \\ &= pn \left[(t+q)^{n-1}\right]_{t=p} \\ &= np\,. \end{aligned}$$

De même

$$\begin{aligned} E(X^2) &= \sum_{k=0}^{n} k^2\, C_n^k\, p^k q^{n-k} \\ &= \sum_{k=0}^{n} (k(k-1)+k)\, C_n^k\, p^k q^{n-k} \\ &= p^2 \sum_{k=2}^{n} k(k-1)\, C_n^k\, p^{k-2} q^{n-k} + E(X) \\ &= p^2 \left[\frac{d^2}{dt^2} \sum_{k=0}^{n} C_n^k\, t^k q^{n-k}\right]_{t=p} + np \\ &= p^2 \left[\frac{d^2}{dt^2}(t+q)^n\right]_{t=p} + np \\ &= p^2 n(n-1) \left[(t+q)^{n-2}\right]_{t=p} + np \\ &= n^2p^2 + np(1-p) \\ &= n^2p^2 + npq \end{aligned}$$

donc

$$\operatorname{var}(X) = E(X^2) - (E(X))^2 = n^2p^2 + npq - (np)^2 = npq\,.$$

**Loi de Poisson $\mathcal{P}(\lambda)$.**

Soit $X$ une v.a. qui suit la loi $\mathcal{P}(\lambda)$, $\lambda > 0$. La série

$$\sum_{k=0}^{+\infty} P(X=k)\,k^2 = \sum_{k=0}^{+\infty} e^{-\lambda}\,\frac{\lambda^k}{k!}\,k^2$$

est convergente d'après la règle de d'Alembert, donc $X$ est de carré sommable. On a alors

$$E(X) = \sum_{k=0}^{+\infty} P(X=k)\,k = \sum_{k=0}^{+\infty} e^{-\lambda}\,\frac{\lambda^k}{k!}\,k = e^{-\lambda}\sum_{k=1}^{+\infty}\frac{\lambda^k}{(k-1)!}$$
$$= \lambda e^{-\lambda}\sum_{k=1}^{+\infty}\frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda}e^{\lambda} = \lambda.$$

$$\begin{aligned}
E(X^2) &= \sum_{k=0}^{+\infty} e^{-\lambda}\,\frac{\lambda^k}{k!}\,k^2 \\
&= \left(\sum_{k=0}^{+\infty} e^{-\lambda}\,k(k-1)\frac{\lambda^k}{k!} + \sum_{k=0}^{+\infty} e^{-\lambda}\,k\frac{\lambda^k}{k!}\right) \\
&= e^{-\lambda}\left(\sum_{k=0}^{+\infty} k(k-1)\frac{\lambda^k}{k!}\right) + E(X) \\
&= e^{-\lambda}\lambda^2\left[\frac{d^2}{dt^2}\,e^t\right]_{t=\lambda} + \lambda \\
&= \lambda^2 + \lambda
\end{aligned}$$

d'où $\operatorname{var}(X) = \lambda$.

**Loi géométrique $\mathcal{G}(p)$.**

Soit $X$ une v.a. qui suit la loi $\mathcal{G}(p)$, $0 < p < 1$. La série

$$\sum_{k=1}^{+\infty} P(X=k)\,k^2 = \sum_{k=1}^{+\infty} k^2 p q^{k-1}$$

est convergente d'après la règle de d'Alembert, donc $X$ est de carré sommable. On a alors

$$E(X) = \sum_{k=1}^{+\infty} P(X=k)\,k = \sum_{k=1}^{+\infty} kpq^{k-1} = p\sum_{k=1}^{+\infty} kq^{k-1}$$
$$= p\left[\frac{d}{dt}\sum_{k=0}^{+\infty} t^k\right]_{t=q} = p\left[\frac{d}{dt}\,\frac{1}{1-t}\right]_{t=q} = \frac{p}{(1-q)^2} = \frac{1}{p}$$

et

$$\begin{aligned}
E(X^2) &= \sum_{k=1}^{+\infty} k^2 p q^{k-1} \\
&= \sum_{k=1}^{+\infty} k(k-1) p q^{k-1} + \sum_{k=1}^{+\infty} k p q^{k-1} \\
&= pq \left[ \frac{d^2}{dt^2} \sum_{k=0}^{+\infty} t^k \right]_{t=q} + E(X) \\
&= pq \left[ \frac{d^2}{dt^2} \frac{1}{1-t} \right]_{t=q} + \frac{1}{p} \\
&= pq \frac{2}{(1-q)^3} + \frac{1}{p} \\
&= \frac{2q}{p^2} + \frac{1}{p} \\
&= \frac{1+q}{p^2}
\end{aligned}$$

d'où

$$\operatorname{var}(X) = E(X^2) - (E(X))^2 = \frac{q}{p^2}.$$

### 8.5.3 Cas d'une loi continue à densité.

Soit $X$ une v.a. continue sur un espace probabilisé $(\Omega, \mathfrak{A}, P)$ dont la loi $P_X$ est définie par une densité de probabilité $g$, *i.e.* une v.a. de densité $g$ sur $\mathbb{R}$ :

$$P(a \leqslant X \leqslant b) = \int_a^b g(t)\, dt \quad \forall a, b \in \mathbb{R},\ a < b.$$

**Espérance.**

On dit que $X$ est *sommable* si l'intégrale impropre

$$\int_{-\infty}^{+\infty} t\, g(t)\, dt \tag{8.74}$$

est absolument convergente. Si $X$ est sommable, l'intégrale (8.74) sera appelée *espérance* de $X$ et notée $E(X)$.

**Moment d'ordre 2.**

**Lemme 8. 11.** *(i) Soit $X$ une v.a. de densité $g$. Alors $X^2$ est une v.a. de densité la fonction $h$ définie par*

$$h(s) = \begin{cases} \frac{g(\sqrt{s}) + g(-\sqrt{s})}{2\sqrt{s}} & si\ s > 0 \\ 0 & si\ s \leqslant 0\,. \end{cases} \tag{8.75}$$

*(ii) $X$ est de carré sommable si seulement si l'intégrale impropre*

$$\int_{-\infty}^{+\infty} t^2\, g(t)\, dt \tag{8.76}$$

*est convergente et l'on a alors*

$$E(X^2) = \int_{-\infty}^{+\infty} t^2\, g(t)\, dt\,. \tag{8.77}$$

**Démonstration.**
(i) On a

$$P(X^2 \leqslant 0) = P(X^2 < 0) + P(X^2 = 0) = P(\varnothing) + P(X = 0) = 0 + \int_0^0 g(t)dt = 0.$$

Soient donc $a, b$ tel que $0 < a < b$. On a

$$P(a \leqslant X^2 \leqslant b) = P(\sqrt{a} \leqslant X \leqslant \sqrt{b}) + P(-\sqrt{b} \leqslant X \leqslant -\sqrt{a}).$$

Or

$$P(\sqrt{a} \leqslant X \leqslant \sqrt{b}) = \int_{\sqrt{a}}^{\sqrt{b}} g(t)\, dt = \int_a^b g(\sqrt{s})\, \frac{ds}{2\sqrt{s}}$$

$$P(-\sqrt{b} \leqslant X \leqslant -\sqrt{a}) = \int_{-\sqrt{b}}^{-\sqrt{a}} g(t)\, dt = \int_{\sqrt{a}}^{\sqrt{b}} g(-t)\, dt = \int_a^b g(-\sqrt{s})\, \frac{ds}{2\sqrt{s}}$$

donc

$$P(a \leqslant X^2 \leqslant b) = \int_a^b g(\sqrt{s})\, \frac{ds}{2\sqrt{s}} + \int_a^b g(-\sqrt{s})\, \frac{ds}{2\sqrt{s}} = \int_a^b \frac{g(\sqrt{s}) + g(-\sqrt{s})}{2\sqrt{s}}\, ds$$

d'où le résultat.
(ii) Par définition, $X$ est de carré sommable si $X^2$ est sommable. D'après (i), cela signifie que l'intégrale impropre

$$\int_0^{+\infty} s\, \frac{g(\sqrt{s}) + g(-\sqrt{s})}{2\sqrt{s}}\, ds \tag{8.78}$$

est convergente. Or pour tous $A > 0$ on a

$$\begin{aligned} \int_0^A s\, \frac{g(\sqrt{s})}{2\sqrt{s}}\, ds &= \int_0^{\sqrt{A}} t^2\, g(t)\, dt \\ \int_0^A s\, \frac{g(-\sqrt{s})}{2\sqrt{s}}\, ds &= \int_{-\sqrt{A}}^0 t^2\, g(t)\, dt \end{aligned}$$

donc l'intégrale (8.78) est convergente si et seulement si l'intégrale (8.76) l'est, et elles sont alors égales. □

On définit le moment d'ordre 2 d'une v.a. de carré sommable par la formule (8.77). La variance et l'écart-type sont alors définis comme dans le cas discret. Tous les résultats énoncés dans le cas discret restent valables dans le cas continu à densité. Nous l'admettrons, car les démonstrations nécessitent l'introduction de la théorie de l'intégration sur l'espace probabilisé $(\Omega, \mathfrak{A}, P)$ sur lequel sont définies les v.a., *i.e.* l'introduction de la notion d'intégrale

$$\int_\Omega X\, dP$$

par rapport à la mesure de probabilité $P$ sur $\Omega$. On montre alors le théorème, dit *de transfert* ([9], Th.1.1., p.130), affirmant que, si $X$ a pour densité $g$,

$$\int_\Omega X\,dP = \int_{-\infty}^{+\infty} t\,g(t)\,dt \tag{8.79}$$

$$\int_\Omega (f\circ X)\,dP = \int_{-\infty}^{+\infty} f(t)\,g(t)\,dt \tag{8.80}$$

avec $f:\mathbb{R}\to\mathbb{R}$ fonction continue quelconque. L'espérance $E(X)$ est alors $\int_\Omega X\,dP$ et le moment d'ordre 2 $E(X^2)$ de $X$ est $\int_\Omega X^2\,dP$. L'équation (8.79) donne immédiatement d'après les propriétés de l'intégrale sur $(\Omega,\mathfrak{A},P)$ la Prop.8.3. Enfin, la formule (8.80) peut s'écrire

$$\int_\Omega (f\circ X)\,dP = \int_{-\infty}^{+\infty} f(t)\,dP_X(t)$$

en notant $dP_X(t)=g(t)dt$, ou encore

$$\int_\Omega (f\circ X)\,dP = \int_{\mathbb{R}} f\,dP_X\,.$$

Sous cette forme, on voit que le formalisme d'intégration par rapport à une mesure de probabilité permet d'inclure le cas discret. Le cas discret correspond en effet au cas où la mesure $P_X$ est une combinaison (8.34) de mesures de Dirac, et l'on a alors

$$E(X)=\int_{\mathbb{R}} t\,dP_X(t)=\sum_{i\in I} P(X=x_i)\,x_i\,,$$

$$E(X^2)=\int_{\mathbb{R}} t^2\,dP_X(t)=\sum_{i\in I} P(X=x_i)\,x_i^2.$$

**Exemples.**

Soit $X$ une v.a. qui suit la loi $\mathcal{N}(0,1)$. On a:

$$E(X)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} t\,e^{-\frac{t^2}{2}}\,dt=0$$

puisque la fonction à intégrer sur $\mathbb{R}$ est impaire.

$$E(X^2)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} t^2\,e^{-\frac{t^2}{2}}\,dt=\frac{1}{\sqrt{2\pi}}\left(\left[-te^{-\frac{t^2}{2}}\right]_{-\infty}^{+\infty}+\int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}}\,dt\right)=1$$

d'où $\mathrm{var}\,(X)=1$ et $\sigma(X)=1$.

Soit $X$ une v.a. qui suit la loi $\mathcal{N}(m,\sigma)$, $m \in \mathbb{R}, \sigma > 0$.

$$\begin{aligned}
E(X) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} t\, e^{-\frac{(t-m)^2}{2\sigma^2}}\, dt \\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (\sigma u + m)\, e^{-\frac{u^2}{2}}\, \sigma\, du \\
&= m \\
E(X^2) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^2\, e^{-\frac{(t-m)^2}{2\sigma^2}}\, dt \\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (\sigma u + m)^2\, e^{-\frac{u^2}{2}}\, \sigma\, du \\
&= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} u^2\, e^{-\frac{u^2}{2}}\, du + \frac{m^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}}\, du \\
&= \sigma^2 + m^2
\end{aligned}$$

d'où $\mathrm{var}\,(X) = \sigma^2$ et $\sigma(X) = \sigma$ .

**Cas de v.a. équidistribuées.**

Soient $X, Y$ des v.a. équidistribuées. Alors $X$ est de carré sommable si et seulement si $Y$ l'est, et alors elles ont même espérance et même écart-type.

### 8.5.4 Inégalité de Bienaymé-Tchebychev.

**Théorème 8. 6.** *Soit $X$ une variable aléatoire de carré sommable. Alors pour tout $t > 0$ :*

$$P\left(|X - E(X)| \geqslant t\right) \quad \leqslant \quad \frac{(\sigma(X))^2}{t^2}. \tag{8.81}$$

**Démonstration.**

Nous faisons la démonstration dans le cas d'une v.a. discrète, le cas d'une v.a. continue à densité étant analogue mais avec des intégrales. Soit $Y = X - E(X)$ la v.a. centrée. Si la loi de $X$ est

$$P_X = \sum_{i \in I} P(X = x_i)\, \delta_{x_i}$$

celle de $Y$ est

$$P_Y = \sum_{i \in I} P(X = x_i)\, \delta_{y_i}$$

avec $y_i = x_i - E(X)$ pour tout $i$. On a

$$\begin{aligned}
E(Y^2) &= \sum_i P(X = x_i)\, y_i^2 \\
&= \sum_{\{i;\, |y_i| \geqslant t\}} P(Y = y_i)\, y_i^2 + \sum_{\{i;\, |y_i| < t\}} P(Y = y_i)\, y_i^2 \\
&\geqslant \sum_{\{i;\, |y_i| \geqslant t\}} P(Y = y_i)\, y_i^2 \\
&\geqslant t^2\, P(|Y| \geqslant t)
\end{aligned}$$

d'où

$$P(|Y| \geqslant t) \leqslant \frac{E(Y^2)}{t^2}. \tag{8.82}$$

Or d'après le Lemme 8.10, $E(Y^2) = \text{var}\,(X)$, donc (8.82) s'écrit

$$P(|X - E(X)| \geqslant t) \leqslant \frac{\text{var}\,(X)}{t^2} = \frac{(\sigma(X))^2}{t^2}.$$

□

D'après l'inégalité de Bienaymé-Tchebychev, on a pour toute v.a. de carré sommable et tout $t > 0$ :

$$P(|X - E(X)| \geqslant t\,\sigma(X)) \leqslant \frac{1}{t^2}$$

d'où

$$P(|X - E(X)| < t\,\sigma(X)) \geqslant 1 - \frac{1}{t^2}. \tag{8.83}$$

En particulier,

$$P(|X - E(X)| < 2\,\sigma(X)) \geqslant 1 - \frac{1}{4} = 0.75$$
$$P(|X - E(X)| < 3\,\sigma(X)) \geqslant 1 - \frac{1}{9} \approx 0.889.$$

Il s'agit de relations *universelles, i.e.* valables pour toute v.a. de carré sommable. Mais elles sont assez grossières. En effet, dans le cas où $X$ suit la loi normale $\mathcal{N}(m, \sigma)$, $Y = \frac{X - E(X)}{\sigma(X)} = \frac{X-m}{\sigma}$ suit la loi normale centrée réduite $\mathcal{N}(0,1)$ et $P(-2 \leqslant Y \leqslant 2) = 2\Phi(2)$ d'après (8.29). Les tables donnent $\Phi(2) \approx 0.4772$ donc $P(-2 \leqslant Y \leqslant 2) \approx 0.954$. De même, $P(-3 \leqslant Y \leqslant 3) \approx 0.997$.

# 8.6 Loi des grands nombres.

## 8.6.1 Loi faible des grands nombres.

Toutes les v.a. considérées dans cette section sont sur un même espace probabilisé $(\Omega, \mathfrak{A}, P)$.

**Théorème 8. 7 (Bernouilli, Tchebychev).** *Soient $(X_n)_{n \geqslant 1}$ une suite de v.a. de carré sommable, indépendantes et équidistribuées. Soient $\mu = E(X_1)$ et $\sigma = \sigma(X_1)$. Pour $n \geqslant 1$, soit $S_n$ la v.a. $X_1 + \ldots + X_n$. Alors pour tout $\varepsilon > 0$,*

$$\lim_{n \to +\infty} P\left(\left\{\left|\frac{S_n}{n} - \mu\right| \geqslant \varepsilon\right\}\right) = 0. \tag{8.84}$$

**Démonstration.**
On a

$$E(\frac{S_n}{n}) = \frac{1}{n}E(S_n) = \frac{1}{n}n\mu = \mu,$$
$$\text{var}\,(\frac{S_n}{n}) = \frac{1}{n^2}\text{var}\,(S_n) = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}.$$

L'inégalité de Bienaymé-Tchebychev (8.81) donne donc

$$0 \leqslant P\left(\left\{\left|\frac{S_n}{n} - \mu\right| \geqslant \varepsilon\right\}\right) \leqslant \frac{\sigma^2}{n\varepsilon^2}$$

donc $P\left(\left\{\left|\frac{S_n}{n} - \mu\right| \geqslant \varepsilon\right\}\right) \to \quad 0$ quand $n \to +\infty$ . □

Dans le Théorème 8.7, la loi faible des grands nombres (8.84) est établie avec l'hypothèse que les $X_n$ sont de carré sommable. Le résultat (8.84) reste vrai sous l'hypothèse plus faible que les $X_n$ sont sommables. Mais cette généralisation requiert des techniques plus avancées ([18], Th.2 page 325).

La loi faible des grands nombres dit que, pour tout $\varepsilon > 0$, on a

$$\lim_{n \to +\infty} P(A_n(\varepsilon)) = 0,$$

en notant $A_n(\varepsilon) = \left\{\left|\frac{S_n}{n} - \mu\right| \geqslant \varepsilon\right\}$ . Mais elle ne *dit pas* que

$$\lim_{n \to +\infty} P\left(\bigcup_{k \geqslant n} A_k(\varepsilon)\right) = 0.$$

En effet, la seule chose que l'on puisse affirmer sur $P\left(\bigcup_{k \geqslant n} A_k(\varepsilon)\right)$ pour l'instant, c'est que $P\left(\bigcup_{k \geqslant n} A_k(\varepsilon)\right) \leqslant \sum_{k \geqslant n} P(A_k(\varepsilon))$ et que chaque terme de la série pris individuellement tend vers 0 quand $n$ tend vers $+\infty$.

En passant aux complémentaires,

$$\lim_{n \to +\infty} P((A_n(\varepsilon))^c) = \lim_{n \to +\infty} P\left(\left\{\left|\frac{S_n}{n} - \mu\right| < \varepsilon\right\}\right) = 1 ,$$

mais on ne sait pas si

$$\lim_{n \to +\infty} P(\bigcap_{k \geqslant n} (A_k(\varepsilon))^c) = \lim_{n \to +\infty} P\left(\left\{\left|\frac{S_k}{k} - \mu\right| < \varepsilon \quad \forall k \geqslant n\right\}\right) = 1 .$$

Autrement dit, on sait que pour tout indice *individuel* $n$ fixé très grand il est "presque sûr" que $\left|\frac{S_n}{n} - \mu\right| < \varepsilon$ mais on ne sait pas s'il est "presque sûr" que $\left|\frac{S_n}{n} - \mu\right| < \varepsilon$ *et* $\left|\frac{S_k}{k} - \mu\right|$ reste $< \varepsilon$ pour tous les $k > n$.

On ne sait donc en particulier pas s'il est "presque sûr" que $\frac{S_n}{n} \to \mu$ quand $n \to +\infty$, *i.e.*

$$P\left(\left\{\frac{S_n}{n} \to \mu\right\}\right) = 1$$

ou encore

$$P\left(\left\{\frac{S_n}{n} \not\to \mu\right\}\right) = 0 .$$

C'est précisément le contenu de la loi forte des grands nombres.

### 8.6.2 Loi forte des grands nombres.

**Théorème 8. 8 (Borel, Kolmogorov).** *Soient $(X_n)_{n\geqslant 1}$ une suite de v.a. sommables, indépendantes et équidistribuées et $\mu = E(X_1)$. Pour $n \geqslant 1$, soit $S_n$ la v.a. $X_1 + \ldots + X_n$ . Alors*

$$P\left(\left\{\frac{S_n}{n} \to \mu\right\}\right) = 1 \tag{8.85}$$

*où $\left\{\frac{S_n}{n} \to \mu\right\}$ désigne l'ensemble des $\omega$ de l'ensemble fondamental $\Omega$ tels que $\frac{S_n(\omega)}{n} \to \mu$ quand $n \to +\infty$.*

Nos admettons aussi ce théorème ([18], Th.3 page 391). Pour une démonstration utilisant très peu de théorie de la Mesure dans le cas de v.a. élémentaires de Bernouilli, voir [20].

### 8.6.3 Convergence en probabilité et convergence presque sûre.

**Définition 8. 30.** *Soit $(X_n)_{n\geqslant 1}$ une suite de v.a. On dit que $X_n$ converge en probabilité (ou en mesure) vers une v.a. $X$ quand $n \to +\infty$ , et l'on note alors $X_n \overset{P}{\to} X$, si la condition suivante est satisfaite :*

$$\forall \varepsilon > 0, \quad \lim_{n\to+\infty} P\left(\{|X_n - X| \geqslant \varepsilon\}\right) = 0 \, .$$

**Définition 8. 31.** *Soit $(X_n)_{n\geqslant 1}$ une suite de v.a. On dit que $X_n$ converge presque sûrement (ou presque partout) vers une v.a. $X$ quand $n \to +\infty$ , et l'on note alors $X_n \to X$ $(P - p.s.)$, ou $X_n \to X$ $(P - p.p.)$, si la condition suivante est satisfaite :*

$$P\left(\{X_n \not\to X\}\right) = 0 \, .$$

Avec ces définitions la loi faible des grands nombres s'énonce :

$$\frac{S_n}{n} \overset{P}{\to} \mu \quad \text{quand } n \to +\infty,$$

et la loi forte des grands nombres s'énonce :

$$\frac{S_n}{n} \to \mu \quad (P - p.s.) \quad \text{quand } n \to +\infty \, .$$

## 8.7 Théorème central limite.

Toutes les v.a. considérées dans cette section sont sur un même espace probabilisé $(\Omega, \mathfrak{A}, P)$.

### 8.7.1 Convergence en loi d'une suite de v.a.

**Définition 8. 32.** *Soit $(X_n)_{n\geqslant 1}$ une suite de v.a. et $F_n$ la fonction de répartition de $X_n$. Soit $X$ une v.a. et $F$ sa fonction de répartition. On dit que $X_n$ converge en loi, ou encore faiblement, vers $X$ quand $n \to +\infty$ , et l'on note alors $X_n \Rightarrow X$, si la condition suivante est satisfaite :*

$$\textit{Pour tout point } x \textit{ où } F \textit{ est continue, } F(x) = \lim_{n\to+\infty} F_n(x).$$

**Théorème 8. 9.** *Soit $(X_n)_{n\geqslant 1}$ une suite de v.a. convergeant en loi vers une v.a. $X$. Soient $F_n$ et $F$ les fonctions de répartition de $X_n$ et $X$ respectivement. On suppose $F$ continue sur $\mathbb{R}$. Alors la suite de fonctions $(F_n)_{n\geqslant 1}$ converge vers $F$ uniformément sur $\mathbb{R}$.*

**Démonstration.**
Il faut montrer que:

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geqslant n_0 \quad \forall x \in \mathbb{R} \quad |F_n(x) - F(x)| < \varepsilon \,. \tag{8.86}$$

Soit $\varepsilon > 0$. Comme $\lim_{x\to-\infty} F(x) = 0$, il existe $a \in \mathbb{R}$ tel que $F(a) < \frac{\varepsilon}{2}$. De même, comme $\lim_{x\to+\infty} F(x) = 1$, il existe $b \in \mathbb{R}$ tel que $F(b) > 1 - \frac{\varepsilon}{2}$. La fonction $F$ étant continue sur l'intervalle compact $[a, b]$, elle est uniformément continue sur $[a, b]$. Il existe donc une subdivision

$$a = x_0 < x_1 < \cdots < x_p < x_{p+1} = b$$

telle que

$$0 \leqslant F(x_{k+1}) - F(x_k) < \frac{\varepsilon}{2} \quad \forall k \quad 0 \leqslant k \leqslant p$$

(on rappelle qu'une fonction de répartition est croissante). Comme $X_n \Rightarrow X$, pour chaque $k = 0, 1, \cdots, p+1$ on a $F(x_k) = \lim_{n\to+\infty} F_n(x_k)$. Donc il existe $N_k \in \mathbb{N}$ tel que pour tout $n \geqslant N_k$, $|F_n(x_k) - F(x_k)| < \frac{\varepsilon}{2}$. Soit $n_0 = \sup_{0\leqslant k\leqslant p+1} N_k$. Alors

$$|F_n(x_k) - F(x_k)| < \frac{\varepsilon}{2} \quad \forall n \geqslant n_0 \,. \tag{8.87}$$

Soit maintenant $x \in \mathbb{R}$ quelconque.
• Si $x \leqslant a$,

$$\begin{aligned} F_n(x) - F(x) &\leqslant F_n(x) \quad (\text{car } F \text{ est } \geqslant 0) \\ &\leqslant F_n(a) \quad (\text{car } F_n \text{ est croissante}) \\ &\leqslant F_n(a) - F(a) + F(a) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \qquad \forall n \geqslant n_0 \,. \end{aligned}$$

Mais

$$F_n(x) - F(x) \geqslant -F(x) \geqslant -F(a) > -\frac{\varepsilon}{2} > -\varepsilon \,.$$

Donc $|F_n(x) - F(x)| < \varepsilon \quad \forall n \geqslant n_0$.
• Si $x \geqslant b$,

$$\begin{aligned} F_n(x) - F(x) &\leqslant 1 - F(x) \quad (\text{car } F_n \leqslant 1) \\ &\leqslant 1 - F(b) \quad (\text{car } F \text{ est croissante}) \\ &\leqslant \frac{\varepsilon}{2} < \varepsilon \,. \end{aligned}$$

Mais

$$F_n(x) - F(x) \geqslant F_n(b) - 1 = F_n(b) - F(b) - (1 - F(b)).$$

Comme $0 \leqslant 1 - F(b) < \frac{\varepsilon}{2}$, on a $-(1 - F(b)) > -\frac{\varepsilon}{2}$.
Or $F_n(b) - F(b) > -\frac{\varepsilon}{2}\ \forall n \geqslant n_0$. Donc

$$F_n(x) - F(x) > -\frac{\varepsilon}{2} - \frac{\varepsilon}{2} = -\varepsilon \quad \forall n \geqslant n_0 \,.$$

D'où $|F_n(x) - F(x)| < \varepsilon \quad \forall n \geqslant n_0$.
- Si $x_k \leqslant x \leqslant x_{k+1}$, $0 \leqslant k \leqslant p$,

$$F_n(x) - F(x) \leqslant F_n(x_{k+1}) - F(x_k) = F_n(x_{k+1}) - F(x_{k+1}) + F(x_{k+1}) - F(x_k).$$

Or $|F_n(x_{k+1}) - F(x_{k+1})| < \frac{\varepsilon}{2}$ pour $n \geqslant n_0$ et $0 \leqslant F(x_{k+1}) - F(x_k) < \frac{\varepsilon}{2}$ par définition de la subdivision. Donc

$$F_n(x) - F(x) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \quad \forall n \geqslant n_0 .$$

Mais

$$F_n(x) - F(x) \geqslant F_n(x_k) - F(x_{k+1}) = F_n(x_k) - F(x_k) + F(x_k) - F(x_{k+1}) .$$

Comme $|F_n(x_k) - F(x_k)| < \frac{\varepsilon}{2} \, \forall n \geqslant n_0$ et $0 \leqslant F(x_{k+1}) - F(x_k) < \frac{\varepsilon}{2}$ ,

$$F_n(x) - F(x) > -\frac{\varepsilon}{2} - \frac{\varepsilon}{2} = -\varepsilon \qquad \forall n \geqslant n_0 .$$

D'où $|F_n(x) - F(x)| < \varepsilon \quad \forall n \geqslant n_0$. (8.86) est démontrée. □

### 8.7.2 Théorème central limite.

**Théorème 8. 10.** *Soient $(X_n)_{n\geqslant 1}$ une suite de v.a. de carré sommable, indépendantes et équidistribuées. Soient $\mu = E(X_1)$ et $\sigma = \sigma(X_1)$. Pour $n \geqslant 1$, soit $S_n$ la v.a. $X_1 + \ldots + X_n$ et $\widetilde{S_n} = \frac{S_n - n\mu}{\sigma\sqrt{n}}$ la variable centrée réduite.*

*Alors $\widetilde{S_n}$ converge en loi quand $n \to +\infty$ vers une v.a. qui a pour loi la loi normale centrée réduite $\mathcal{N}(0,1)$, i.e.*

$$\lim_{n\to+\infty} P(\widetilde{S_n} \leqslant t) = \Psi(t) \quad \forall t \in \mathbb{R} . \tag{8.88}$$

*On notera simplement $\widetilde{S_n} \Rightarrow \mathcal{N}(0,1)$ quand $n \to +\infty$.*

Nous admettons ce théorème dont la démonstration nécessiterait l'introduction de notions dépassant le cadre de ce cours, telles que fonctions caractéristiques de v.a. et Théorème de la convergence dominée de Lebesgue ([18], Th.3 page 326).

### 8.7.3 Théorème de Berry-Esseen.

**Théorème 8. 11.** *Dans les conditions du Théorème central limite, notons $\widetilde{F_n}$ la fonction de répartition de $\widetilde{S_n}$.*

*On suppose de plus que $\left|\widetilde{X_1}\right|^3$ est sommable, où $\widetilde{X_1} = \frac{X_1 - \mu}{\sigma}$. Alors*

$$\left|\widetilde{F_n}(t) - \Psi(t)\right| \leqslant 0.8 \, \frac{E\left(\left|\widetilde{X_1}\right|^3\right)}{\sqrt{n}} \quad \forall t \in \mathbb{R} . \tag{8.89}$$

Nous admettons aussi ce théorème ([18], page 374).

### 8.7.4 Applications à la loi binomiale.

**Application du Théorème central limite.**

Le Théorème central limite s'applique en particulier au cas où les $X_n$ sont des v.a. élémentaires de Bernouilli indépendantes et équidistribuées, d'espérance $\mu = p$ $(0 < p < 1)$ et écart-type $\sigma = \sqrt{pq}$.

Alors $S_n$ suit la loi binomiale $\mathcal{B}(n,p)$ et $\widetilde{S_n} \Rightarrow \mathcal{N}(0,1)$ (une démonstration directe de ce fait est aussi donnée dans la section 8.8 Appendice 1).

**Application du Théorème de Berry-Essen.**

Comme la v.a. $|X_1 - p|^3$ ne prend que les 2 valeurs $p^3$ et $q^3$ avec les probabilités respectives $q$ et $p$, elle est sommable et $E(|X_1 - p|^3) = qp^3 + pq^3 = pq(p^2+q^2)$. Donc $\left|\widetilde{X_1}\right|^3$ est sommable et

$$E\left(\left|\widetilde{X_1}\right|^3\right) = E\left(\frac{|X_1 - p|^3}{\sigma^3}\right) = \frac{1}{\sigma^3} E\left(|X_1 - p|^3\right) = \frac{p^2+q^2}{\sqrt{pq}}.$$

Le Théorème de Berry-Esseen s'applique donc, et montre que si $F_n$ désigne la fonction de répartition d'une v.a. $S_n$ suivant la loi binomiale $\mathcal{B}(n,p)$ et si $\widetilde{F_n}$ désigne la fonction de répartition de la v.a. centrée réduite $\widetilde{S_n} = \frac{S_n - np}{\sqrt{npq}}$, alors

$$\left|\widetilde{F_n}(t) - \Psi(t)\right| \leqslant \gamma_n(p) \quad \forall t \in \mathbb{R} \tag{8.90}$$

avec

$$\gamma_n(p) = 0.8\, \frac{p^2+q^2}{\sqrt{npq}}\,. \tag{8.91}$$

Or

$$\widetilde{F_n}\Big(\frac{x-np}{\sqrt{npq}}\Big) = F_n(x) \quad \forall x \in \mathbb{R}\,.$$

Donc (8.90) s'écrit encore en posant $t = \frac{x-np}{\sqrt{npq}}$

$$\left|F_n(x) - \Psi\Big(\frac{x-np}{\sqrt{npq}}\Big)\right| \leqslant \gamma_n(p) \quad \forall x \in \mathbb{R}\,. \tag{8.92}$$

**Approximation normale.**

Si $S_n$ est une v.a. suivant la loi binomiale $\mathcal{B}(n,p)$, on peut ainsi approcher $\widetilde{S_n} = \frac{S_n - np}{\sqrt{npq}}$ par la loi normale centrée réduite $\mathcal{N}(0,1)$, et donc $S_n$ par $\mathcal{N}(np, \sqrt{npq})$.

On dit que $\mathcal{N}(np, \sqrt{npq})$ est *l'approximation normale* de la loi binomiale $\mathcal{B}(n,p)$.

Avec l'approximation normale, on a pour $\alpha, \beta, a, b \in \mathbb{R}$

$$P(\alpha < \widetilde{S_n} \leqslant \beta) \;\approx\; \Psi(\beta) - \Psi(\alpha) \tag{8.93}$$

$$P(a < S_n \leqslant b) \;\approx\; \Psi\Big(\frac{b-np}{\sqrt{npq}}\Big) - \Psi\Big(\frac{a-np}{\sqrt{npq}}\Big)\,. \tag{8.94}$$

D'après (8.91), cette approximation peut être insuffisante si $npq$ est petit, en particulier même si $n$ est très grand mais $p$ très petit. Une règle empirique souvent adoptée est que l'approximation normale est acceptable pour

$$n \inf(p,q) \geqslant 6\,. \tag{8.95}$$

**Approximation normale avec correction de continuité.**

Posons pour $0 \leqslant k \leqslant n$

$$t_k = \frac{k-np}{\sqrt{npq}}\,. \tag{8.96}$$

On a

$$P(\widetilde{S_n} = t_k) = P(S_n = k) = C_n^k\, p^k q^{n-k}\,.$$

On veut comparer

$$P(\widetilde{S_n} = t_k) = C_n^k\, p^k q^{n-k} \tag{8.97}$$

et

$$\frac{1}{\sqrt{2\pi}}\, e^{-\frac{t_k^2}{2}}\,. \tag{8.98}$$

Si $k$ est fixé, on a $C_n^k = \frac{n(n-1)\cdots(n-k+1)}{k!} \leqslant \frac{n^k}{k!}$ donc

$$C_n^k\, p^k q^{n-k} \leqslant \frac{n^k}{k!} p^k q^{n-k} \to 0$$

quand $n \to +\infty$ puisque $0 < q < 1$. D'autre part, $\lim_{n\to+\infty} t_k = -\infty$, donc les deux quantités (8.97) et (8.98) sont négligeables.

Pour avoir des quantités qui ne soient pas négligeables, on est amené à supposer que $t_k$ reste borné, *i.e.* que $k$ est en fait une fonction $k(n)$ qui tend vers $+\infty$ avec $n$ et telle qu'il existe un $T > 0$ pour lequel

$$|t_{k(n)}| \leqslant T \quad \forall\, n \in \mathbb{N}^*\,. \tag{8.99}$$

On notera simplement $k$ au lieu de $k(n)$. Alors la condition (8.99) implique, d'après le Lemme 8.13 de l'Appendice 1, que

$$C_n^k\, p^k q^{n-k} \sim \frac{1}{\sqrt{npq}} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{t_k^2}{2}} \quad \text{quand } n \to +\infty\,. \tag{8.100}$$

On notera que si $k$ *est fixé* (8.100) n'est pas nécessairement vérifiée : par exemple pour $k = 0$, $t_0 = -\sqrt{\frac{np}{q}}$ et

$$q^n \not\sim \frac{1}{\sqrt{npq}} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{t_0^2}{2}} = \frac{1}{\sqrt{npq}} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{np}{2q}}\,.$$

En effet,

$$\frac{1}{\sqrt{npq}} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{np}{2q} - n\mathrm{Log}\, q} = \frac{1}{\sqrt{2\pi pq}} e^{-\frac{n}{2q}(1-q+2q\mathrm{Log}\, q) - \frac{1}{2}\mathrm{Log}\, n}$$

tend vers $+\infty$ (*resp.* 0) quand $n \to +\infty$ si $1-q+2q\mathrm{Log}\, q < 0$ (*resp.* $\geqslant 0$) (on vérifie immédiatement que la fonction $t \mapsto 1-t+2t\mathrm{Log}\, t$ sur l'intervalle $]0,1[$ possède un unique zéro).

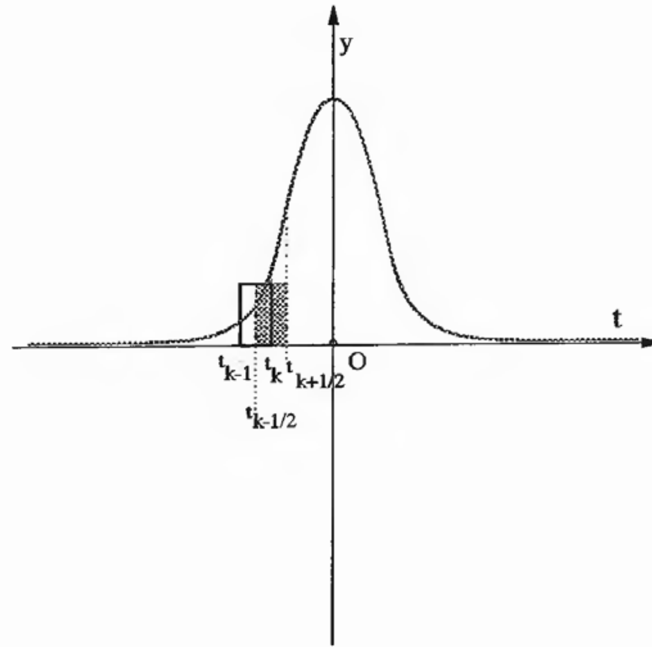



FIG. 8.1: *Correction de continuité.*

Considérons maintenant l'approximation normale de $\widetilde{F_n}$ par $\Psi$. Elle approche

$$P(S_n = k) = C_n^k\, p^k q^{n-k} = \widetilde{F_n}(t_k) - \widetilde{F_n}(t_{k-1})$$

par

$$\Psi(t_k) - \Psi(t_{k-1}) = \frac{1}{\sqrt{2\pi}} \int_{t_{k-1}}^{t_k} e^{-\frac{t^2}{2}}\, dt\,.$$

Or comme $t_k - t_{k-1} = \frac{1}{\sqrt{npq}}$ , $C_n^k\, p^k q^{n-k}$ est l'aire du rectangle

$$\mathcal{R}_k = [t_{k-1}, t_k] \times [0, \sqrt{npq}\, C_n^k\, p^k q^{n-k}].$$

D'après (8.100), l'approximation normale consiste donc à approcher l'aire $\widetilde{F_n}(t_k) - \widetilde{F_n}(t_{k-1})$ de $\mathcal{R}_k$ par l'aire $\Psi(t_k) - \Psi(t_{k-1})$ sous la courbe en cloche $y = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$ comprise entre $t_{k-1}$ et $t_k$. On conçoit graphiquement (voir Fig. 8.1) que cette approximation pourrait être meilleure si l'on approchait l'aire de $\mathcal{R}_k$ par l'aire sous la courbe en cloche comprise entre les valeurs $t_{k-\frac{1}{2}} = \frac{k-\frac{1}{2}-np}{\sqrt{npq}} = t_{k-1} + \frac{1}{2\sqrt{npq}}$ et $t_{k+\frac{1}{2}} = \frac{k+\frac{1}{2}-np}{\sqrt{npq}} = t_k + \frac{1}{2\sqrt{npq}}$. En effet, des compensations entre "aire au dessus" et "aire au dessous" apparaissent dans ce cas. Cette nouvelle aire est

$$\Psi(t_k + \frac{1}{2\sqrt{npq}}) - \Psi(t_{k-1} + \frac{1}{2\sqrt{npq}}).$$

Introduisons alors la fonction $\Psi^*$ définie par

$$\Psi^*(x) = \Psi(x + \frac{1}{2\sqrt{npq}}) \quad \forall\, x \in \mathbb{R}\,. \tag{8.101}$$

La nouvelle aire est alors $\Psi^*(t_k) - \Psi^*(t_{k-1})$. Approcher l'aire de $\mathcal{R}_k$ par la nouvelle aire revient donc à approcher $\widetilde{F_n}(x)$ non pas par $\Psi(x)$ mais par $\Psi^*(x)$. Cette approximation est appelée *l'approximation normale avec correction de continuité.* Elle

est effectivement meilleure *en général* pour les valeurs (8.96) de discontinuité de $\widetilde{F_n}$ "proche" de 0, *i.e.* pour les valeurs de $k$ "proches" de $np$, comme on le voit sur les exemples d'approximation ci-dessous. Si $np$ est petit, cela recouvre pratiquement l'ensemble des valeurs de $k$. Mais lorsque $k$ n'est pas "proche" de $np$, l'approximation avec correction de continuité peut être plus mauvaise. Avec l'approximation normale avec correction de continuité , on a pour $\alpha, \beta, a, b \in \mathbb{R}$

$$P(\alpha < \widetilde{S_n} \leqslant \beta) \approx \Psi(\beta + \frac{1}{2\sqrt{npq}}) - \Psi(\alpha + \frac{1}{2\sqrt{npq}}) \qquad (8.102)$$

$$P(a < S_n \leqslant b) \approx \Psi(\frac{b + \frac{1}{2} - np}{\sqrt{npq}}) - \Psi(\frac{a + \frac{1}{2} - np}{\sqrt{npq}}) \, . \qquad (8.103)$$

Les tableaux qui suivent, et les figures 8.2 à 8.4, donnent des exemples de la précision des approximations normales sans et avec correction de continuité.

**Exemples.**

Exemples de valeurs de la fonction de répartition $F$ de $\mathcal{B}(n,p)$ :
$F(k), \Psi(\frac{k-np}{\sqrt{npq}})$ et $\Psi^*(\frac{k-np}{\sqrt{npq}}) = \Psi(\frac{k+\frac{1}{2}-np}{\sqrt{npq}})$

$n = 10$

$p = 0.2 \quad , \quad np = 2 \quad , \quad \gamma_{10}(0.2) = 0.43007 \, .$

| $k$ | $F$ | $\Psi$ | $\Psi^*$ | $F - \Psi$ | $F - \Psi^*$ |
|---|---|---|---|---|---|
| 0 | 0.107374 | 0.056923 | 0.11784 | 0.050451 | −0.010465 |
| 1 | 0.37581 | 0.214598 | 0.346316 | 0.161212 | 0.029493 |
| **2** | 0.6778 | 0.5 | 0.653684 | 0.1778 | 0.024116 |
| 3 | 0.879126 | 0.785402 | 0.88216 | 0.093724 | −0.003033 |
| 4 | 0.967207 | 0.943077 | 0.975947 | 0.02413 | −0.00874 |
| 5 | 0.993631 | 0.991147 | 0.997171 | 0.002484 | −0.00354 |
| 6 | 0.999136 | 0.999217 | 0.999813 | −0.000081 | −0.000677 |
| 7 | 0.999922 | 0.999961 | 0.999993 | −0.000039 | −0.000071 |
| 8 | 0.999996 | 0.999999 | 1.0 | −0.000003 | −0.000004 |
| 9 | 1.0 | 1.0 | 1.0 | 0 | 0 |

$$p = 0.5 \quad , \quad np = 5 \quad , \quad \gamma_{10}(0.5) = 0.252982 \ .$$

| $k$ | $F$ | $\Psi$ | $\Psi^*$ | $F - \Psi$ | $F - \Psi^*$ |
|---|---|---|---|---|---|
| 0 | 0.000977 | 0.000783 | 0.002213 | 0.000194 | −0.001236 |
| 1 | 0.010742 | 0.005706 | 0.013428 | 0.005036 | −0.002686 |
| 2 | 0.054688 | 0.02889 | 0.056923 | 0.025798 | −0.002235 |
| 3 | 0.171875 | 0.102952 | 0.171391 | 0.068923 | 0.000484 |
| 4 | 0.376953 | 0.263545 | 0.375915 | 0.113408 | 0.001038 |
| 5 | 0.623047 | 0.5 | 0.624085 | 0.123047 | −0.001038 |
| 6 | 0.828125 | 0.736455 | 0.828609 | 0.09167 | −0.000484 |
| 7 | 0.945313 | 0.897048 | 0.943077 | 0.048264 | 0.002236 |
| 8 | 0.989258 | 0.97111 | 0.986572 | 0.018148 | 0.002686 |
| 9 | 0.999023 | 0.994294 | 0.997787 | 0.004729 | 0.001237 |
| 10 | 1.0 | 0.999217 | 0.999748 | 0.000783 | 0.000252 |

$$n = 25$$

$$p = 0.2 \quad , \quad np = 5 \quad , \quad \gamma_{25}(0.2) = 0.272 \ .$$

| $k$ | $F$ | $\Psi$ | $\Psi^*$ | $F - \Psi$ | $F - \Psi^*$ |
|---|---|---|---|---|---|
| 0 | 0.003778 | 0.00621 | 0.012224 | −0.002431 | −0.008446 |
| 1 | 0.02739 | 0.02275 | 0.040059 | 0.00464 | −0.012669 |
| 2 | 0.098225 | 0.066807 | 0.10565 | 0.031418 | −0.007424 |
| 3 | 0.233993 | 0.158655 | 0.226627 | 0.075338 | 0.007366 |
| 4 | 0.420674 | 0.308538 | 0.401294 | 0.112137 | 0.019381 |
| 5 | 0.616689 | 0.5 | 0.598706 | 0.116689 | 0.017983 |
| 6 | 0.780035 | 0.691462 | 0.773373 | 0.088573 | 0.006663 |
| 7 | 0.890877 | 0.841345 | 0.89435 | 0.049532 | −0.003473 |
| 8 | 0.953226 | 0.933193 | 0.959941 | 0.020033 | −0.006715 |
| 9 | 0.982668 | 0.97725 | 0.987776 | 0.005418 | −0.005107 |
| 10 | 0.994445 | 0.99379 | 0.99702 | 0.000655 | −0.002575 |
| 11 | 0.99846 | 0.99865 | 0.999423 | −0.00019 | −0.000963 |
| 12 | 0.999631 | 0.999767 | 0.999912 | −0.000136 | −0.00028 |
| 13 | 0.999924 | 0.999968 | 0.999989 | −0.000044 | −0.000065 |
| 14 | 0.999986 | 0.999997 | 0.999999 | −0.00001 | −0.000012 |
| 15 | 0.999998 | 1.0 | 1.0 | −0.000001 | −0.000001 |
| 16 | 1.0 | 1.0 | 1.0 | 0 | 0 |

$$p = 0.5 \quad , \quad np = 12.5 \quad , \quad \gamma_{25}(0.5) = 0.16 \ .$$

| $k$ | $F$ | $\Psi$ | $\Psi^*$ | $F - \Psi$ | $F - \Psi^*$ |
|---|---|---|---|---|---|
| 0 | 0 | 0 | 0.000001 | 0 | 0 |
| 1 | 0.000001 | 0.000002 | 0.000005 | −0.000001 | −0.000004 |
| 2 | 0.00001 | 0.000013 | 0.000032 | −0.000003 | −0.000021 |
| 3 | 0.000078 | 0.000072 | 0.000159 | 0.000006 | −0.00008 |
| 4 | 0.000455 | 0.000337 | 0.000687 | 0.000118 | −0.000231 |
| 5 | 0.002039 | 0.00135 | 0.002555 | 0.000689 | −0.000516 |
| 6 | 0.007317 | 0.004661 | 0.008198 | 0.002655 | −0.00088 |
| 7 | 0.021643 | 0.013903 | 0.02275 | 0.007739 | −0.001107 |
| 8 | 0.053876 | 0.03593 | 0.054799 | 0.017946 | −0.000923 |
| 9 | 0.114761 | 0.080757 | 0.11507 | 0.034005 | −0.000308 |
| 10 | 0.212178 | 0.158655 | 0.211855 | 0.053523 | 0.000323 |
| 11 | 0.345019 | 0.274253 | 0.344578 | 0.070766 | 0.000441 |
| **12** | 0.5 | 0.42074 | 0.5 | 0.07926 | 0 |
| **13** | 0.654981 | 0.57926 | 0.655422 | 0.075721 | −0.00044 |
| 14 | 0.787822 | 0.725747 | 0.788145 | 0.062075 | −0.000322 |
| 15 | 0.885239 | 0.841345 | 0.88493 | 0.043894 | 0.000308 |
| 16 | 0.946124 | 0.919243 | 0.945201 | 0.026881 | 0.000923 |
| 17 | 0.978357 | 0.96407 | 0.97725 | 0.014288 | 0.001108 |
| 18 | 0.992683 | 0.986097 | 0.991802 | 0.006587 | 0.000881 |
| 19 | 0.997961 | 0.995339 | 0.997445 | 0.002623 | 0.000516 |
| 20 | 0.999545 | 0.99865 | 0.999313 | 0.000895 | 0.000232 |
| 21 | 0.999922 | 0.999663 | 0.999841 | 0.000259 | 0.000081 |
| 22 | 0.99999 | 0.999928 | 0.999968 | 0.000063 | 0.000022 |
| 23 | 0.999999 | 0.999987 | 0.999995 | 0.000013 | 0.000005 |
| 24 | 1.0 | 0.999998 | 0.999999 | 0.000002 | 0.000001 |
| 25 | 1.0 | 1.0 | 1.0 | 0 | 0 |

$n = 50$

$p = 0.2 \quad , \quad np = 10 \quad , \quad \gamma_{50}(0.2) = 0.192333 \ .$

| $k$ | $F$ | $\Psi$ | $\Psi^*$ | $F - \Psi$ | $F - \Psi^*$ |
|---|---|---|---|---|---|
| 0 | 0.000014 | 0.000203 | 0.000391 | −0.000189 | −0.000377 |
| 1 | 0.000193 | 0.000731 | 0.001327 | −0.000538 | −0.001134 |
| 2 | 0.001285 | 0.002339 | 0.004005 | −0.001053 | −0.002719 |
| 3 | 0.005656 | 0.006664 | 0.010778 | −0.001007 | −0.005121 |
| 4 | 0.018496 | 0.016947 | 0.025915 | 0.001549 | −0.007418 |
| 5 | 0.048027 | 0.03855 | 0.055806 | 0.009477 | −0.007778 |
| 6 | 0.103398 | 0.07865 | 0.107962 | 0.024749 | −0.004564 |
| 7 | 0.19041 | 0.144422 | 0.18838 | 0.045988 | 0.00203 |
| 8 | 0.307332 | 0.23975 | 0.297942 | 0.067582 | 0.00939 |
| 9 | 0.44374 | 0.361837 | 0.429842 | 0.081904 | 0.013899 |
| **10** | 0.583559 | 0.5 | 0.570158 | 0.083559 | 0.013401 |
| 11 | 0.710668 | 0.638163 | 0.702058 | 0.072504 | 0.008609 |
| 12 | 0.813943 | 0.76025 | 0.81162 | 0.053693 | 0.002323 |
| 13 | 0.889413 | 0.855578 | 0.892038 | 0.033836 | −0.002624 |
| 14 | 0.939278 | 0.92135 | 0.944194 | 0.017928 | −0.004916 |
| 15 | 0.969197 | 0.96145 | 0.974085 | 0.007747 | −0.004888 |
| 16 | 0.985558 | 0.983053 | 0.989222 | 0.002506 | −0.003663 |
| 17 | 0.993739 | 0.993336 | 0.995995 | 0.000403 | −0.002255 |
| 18 | 0.997489 | 0.997661 | 0.998673 | −0.000172 | −0.001184 |
| 19 | 0.999068 | 0.999269 | 0.999609 | −0.000201 | −0.00054 |
| 20 | 0.999679 | 0.999797 | 0.999897 | −0.000117 | −0.000217 |
| 21 | 0.999898 | 0.99995 | 0.999976 | −0.000051 | −0.000078 |
| 22 | 0.99997 | 0.999989 | 0.999995 | −0.000019 | −0.000025 |
| 23 | 0.999992 | 0.999998 | 0.999999 | −0.000006 | −0.000007 |
| 24 | 0.999998 | 1.0 | 1.0 | −0.000001 | −0.000001 |
| 25 | 1.0 | 1.0 | 1.0 | 0 | 0 |

$$p = 0.5 \quad , \quad np = 25 \quad , \quad \gamma_{50}(0.5) = 0.113137 \ .$$

| $k$ | $F$ | $\Psi$ | $\Psi^*$ | $F - \Psi$ | $F - \Psi^*$ |
|---|---|---|---|---|---|
| 8 | 0.000001 | 0.000001 | 0.000002 | 0 | 0 |
| 9 | 0.000003 | 0.000003 | 0.000006 | 0 | −0.000003 |
| 10 | 0.000012 | 0.000011 | 0.000021 | 0.000001 | −0.000008 |
| 11 | 0.000045 | 0.000038 | 0.000067 | 0.000008 | −0.000022 |
| 12 | 0.000153 | 0.000118 | 0.000203 | 0.000035 | −0.00005 |
| 13 | 0.000468 | 0.000344 | 0.000572 | 0.000124 | −0.000103 |
| 14 | 0.001301 | 0.000931 | 0.00149 | 0.00037 | −0.000188 |
| 15 | 0.0033 | 0.002339 | 0.003605 | 0.000961 | −0.000304 |
| 16 | 0.007673 | 0.005455 | 0.008105 | 0.002219 | −0.000431 |
| 17 | 0.01642 | 0.011826 | 0.016947 | 0.004594 | −0.000527 |
| 18 | 0.032454 | 0.023857 | 0.032996 | 0.008597 | −0.000541 |
| 19 | 0.05946 | 0.044843 | 0.059897 | 0.014617 | −0.000437 |
| 20 | 0.101319 | 0.07865 | 0.101546 | 0.02267 | −0.000226 |
| 21 | 0.161118 | 0.12895 | 0.161099 | 0.032169 | 0.000019 |
| 22 | 0.239944 | 0.198072 | 0.23975 | 0.041872 | 0.000194 |
| 23 | 0.335906 | 0.285804 | 0.335687 | 0.050102 | 0.000219 |
| 24 | 0.443862 | 0.388649 | 0.443769 | 0.055214 | 0.000094 |
| **25** | 0.556138 | 0.5 | 0.556231 | 0.056138 | −0.000093 |
| 26 | 0.664094 | 0.611351 | 0.664313 | 0.052743 | −0.000218 |
| 27 | 0.760056 | 0.714196 | 0.76025 | 0.04586 | −0.000193 |
| 28 | 0.838882 | 0.801928 | 0.838901 | 0.036954 | −0.000018 |
| 29 | 0.898681 | 0.87105 | 0.898454 | 0.02763 | 0.000227 |
| 30 | 0.94054 | 0.92135 | 0.940103 | 0.019189 | 0.000437 |
| 31 | 0.967546 | 0.955157 | 0.967004 | 0.012389 | 0.000542 |
| 32 | 0.98358 | 0.976143 | 0.983053 | 0.007438 | 0.000528 |
| 33 | 0.992327 | 0.988174 | 0.991895 | 0.004152 | 0.000431 |
| 34 | 0.9967 | 0.994545 | 0.996395 | 0.002155 | 0.000305 |
| 35 | 0.998699 | 0.997661 | 0.99851 | 0.001038 | 0.000189 |
| 36 | 0.999532 | 0.999069 | 0.999428 | 0.000463 | 0.000103 |
| 37 | 0.999847 | 0.999656 | 0.999797 | 0.000191 | 0.000051 |
| 38 | 0.999955 | 0.999882 | 0.999933 | 0.000073 | 0.000022 |
| 39 | 0.999988 | 0.999962 | 0.999979 | 0.000026 | 0.000009 |
| 40 | 0.999997 | 0.999989 | 0.999994 | 0.000008 | 0.000003 |
| 41 | 0.999999 | 0.999997 | 0.999998 | 0.000002 | 0.000001 |
| 42 | 1.0 | 0.999999 | 1.0 | 0.000001 | 0 |
| 43 | 1.0 | 1.0 | 1.0 | 0 | 0 |

$$n = 100$$

$$p = 0.2 \quad , \quad np = 20 \quad , \quad \gamma_{100}(0.2) = 0.136 \ .$$

| $k$ | $F$ | $\Psi$ | $\Psi^*$ | $F - \Psi$ | $F - \Psi^*$ |
|---|---|---|---|---|---|
| 0 | 0 | 0 | 0.000001 | 0 | 0 |
| 1 | 0 | 0.000001 | 0.000002 | −0.000001 | −0.000001 |
| 2 | 0 | 0.000003 | 0.000006 | −0.000003 | −0.000006 |
| 3 | 0.000001 | 0.000011 | 0.000019 | −0.00001 | −0.000017 |
| 4 | 0.000004 | 0.000032 | 0.000053 | −0.000027 | −0.000049 |
| 5 | 0.000019 | 0.000088 | 0.000144 | −0.000069 | −0.000125 |
| 6 | 0.000078 | 0.000233 | 0.000369 | −0.000154 | −0.000291 |
| 7 | 0.000277 | 0.000577 | 0.000889 | −0.0003 | −0.000612 |
| 8 | 0.000855 | 0.00135 | 0.00202 | −0.000494 | −0.001164 |
| 9 | 0.002334 | 0.00298 | 0.004332 | −0.000646 | −0.001998 |
| 10 | 0.005696 | 0.00621 | 0.008774 | −0.000513 | −0.003078 |
| 11 | 0.012575 | 0.012224 | 0.016793 | 0.00035 | −0.004218 |
| 12 | 0.025329 | 0.02275 | 0.030396 | 0.002579 | −0.005067 |
| 13 | 0.046912 | 0.040059 | 0.052081 | 0.006853 | −0.005169 |
| 14 | 0.080444 | 0.066807 | 0.084566 | 0.013637 | −0.004122 |
| 15 | 0.128506 | 0.10565 | 0.130295 | 0.022856 | −0.001789 |
| 16 | 0.192338 | 0.158655 | 0.190787 | 0.033682 | 0.001551 |
| 17 | 0.271189 | 0.226627 | 0.265986 | 0.044562 | 0.005203 |
| 18 | 0.362087 | 0.308538 | 0.35383 | 0.05355 | 0.008257 |
| 19 | 0.460161 | 0.401294 | 0.450262 | 0.058868 | 0.0099 |
| **20** | 0.559462 | 0.5 | 0.549738 | 0.059462 | 0.009723 |
| 21 | 0.654033 | 0.598706 | 0.64617 | 0.055327 | 0.007863 |
| 22 | 0.738933 | 0.691462 | 0.734014 | 0.04747 | 0.004918 |
| 23 | 0.810913 | 0.773373 | 0.809213 | 0.03754 | 0.0017 |
| 24 | 0.868647 | 0.841345 | 0.869705 | 0.027302 | −0.001058 |
| 25 | 0.912525 | 0.89435 | 0.915434 | 0.018174 | −0.002909 |
| 26 | 0.944167 | 0.933193 | 0.947919 | 0.010974 | −0.003751 |
| 27 | 0.965848 | 0.959941 | 0.969604 | 0.005908 | −0.003755 |
| 28 | 0.97998 | 0.97725 | 0.983207 | 0.00273 | −0.003226 |
| 29 | 0.988751 | 0.987776 | 0.991226 | 0.000975 | −0.002474 |
| 30 | 0.993941 | 0.99379 | 0.995668 | 0.00015 | −0.001726 |
| 31 | 0.99687 | 0.99702 | 0.99798 | −0.000149 | −0.001109 |
| 32 | 0.99845 | 0.99865 | 0.999111 | −0.0002 | −0.000661 |
| 33 | 0.999263 | 0.999423 | 0.999631 | −0.000159 | −0.000367 |
| 34 | 0.999664 | 0.999767 | 0.999856 | −0.000103 | −0.000191 |
| 35 | 0.999853 | 0.999912 | 0.999947 | −0.000058 | −0.000093 |
| 36 | 0.999938 | 0.999968 | 0.999981 | −0.00003 | −0.000043 |
| 37 | 0.999975 | 0.999989 | 0.999994 | −0.000014 | −0.000018 |
| 38 | 0.99999 | 0.999997 | 0.999998 | −0.000006 | −0.000007 |
| 39 | 0.999996 | 0.999999 | 0.999999 | −0.000002 | −0.000003 |
| 40 | 0.999999 | 1.0 | 1.0 | −0.000001 | −0.000001 |
| 41 | 1.0 | 1.0 | 1.0 | 0 | 0 |

$$p = 0.5 \quad , \quad np = 50 \quad , \quad \gamma_{100}(0.5) = 0.08 \; .$$

| $k$ | $F$ | $\Psi$ | $\Psi^*$ | $F-\Psi$ | $F-\Psi^*$ |
|---|---|---|---|---|---|
| 26 | 0.000001 | 0.000001 | 0.000001 | 0 | 0 |
| 27 | 0.000002 | 0.000002 | 0.000003 | 0 | −0.000001 |
| 28 | 0.000006 | 0.000005 | 0.000009 | 0.000001 | −0.000002 |
| 29 | 0.000016 | 0.000013 | 0.000021 | 0.000003 | −0.000004 |
| 30 | 0.000039 | 0.000032 | 0.000048 | 0.000008 | −0.000008 |
| 31 | 0.000092 | 0.000072 | 0.000108 | 0.000019 | −0.000016 |
| 32 | 0.000204 | 0.000159 | 0.000233 | 0.000045 | −0.000028 |
| 33 | 0.000437 | 0.000337 | 0.000483 | 0.0001 | −0.000046 |
| 34 | 0.000895 | 0.000687 | 0.000968 | 0.000208 | −0.000072 |
| 35 | 0.001759 | 0.00135 | 0.001866 | 0.000409 | −0.000106 |
| 36 | 0.003319 | 0.002555 | 0.003467 | 0.000763 | −0.000148 |
| 37 | 0.006016 | 0.004661 | 0.00621 | 0.001355 | −0.000193 |
| 38 | 0.010489 | 0.008198 | 0.010724 | 0.002292 | −0.000234 |
| 39 | 0.0176 | 0.013903 | 0.017864 | 0.003697 | −0.000264 |
| 40 | 0.028444 | 0.02275 | 0.028717 | 0.005694 | −0.000272 |
| 41 | 0.044313 | 0.03593 | 0.044565 | 0.008383 | −0.000252 |
| 42 | 0.066605 | 0.054799 | 0.066807 | 0.011806 | −0.000201 |
| 43 | 0.096674 | 0.080757 | 0.0968 | 0.015917 | −0.000126 |
| 44 | 0.135627 | 0.11507 | 0.135666 | 0.020557 | −0.000039 |
| 45 | 0.184101 | 0.158655 | 0.18406 | 0.025446 | 0.000041 |
| 46 | 0.242059 | 0.211855 | 0.241964 | 0.030204 | 0.000096 |
| 47 | 0.30865 | 0.274253 | 0.308538 | 0.034397 | 0.000112 |
| 48 | 0.382177 | 0.344578 | 0.382089 | 0.037598 | 0.000088 |
| 49 | 0.460205 | 0.42074 | 0.460172 | 0.039465 | 0.000033 |
| 50 | 0.539795 | 0.5 | 0.539828 | 0.039795 | −0.000033 |
| 51 | 0.617823 | 0.57926 | 0.617911 | 0.038564 | −0.000088 |
| 52 | 0.69135 | 0.655422 | 0.691462 | 0.035929 | −0.000112 |
| 53 | 0.757941 | 0.725747 | 0.758036 | 0.032194 | −0.000095 |
| 54 | 0.815899 | 0.788145 | 0.81594 | 0.027755 | −0.00004 |
| 55 | 0.864373 | 0.841345 | 0.864334 | 0.023029 | 0.00004 |
| 56 | 0.903326 | 0.88493 | 0.9032 | 0.018396 | 0.000127 |
| 57 | 0.933395 | 0.919243 | 0.933193 | 0.014151 | 0.000202 |
| 58 | 0.955687 | 0.945201 | 0.955435 | 0.010486 | 0.000252 |
| 59 | 0.971556 | 0.96407 | 0.971283 | 0.007486 | 0.000273 |
| 60 | 0.9824 | 0.97725 | 0.982136 | 0.00515 | 0.000264 |
| 61 | 0.989511 | 0.986097 | 0.989276 | 0.003414 | 0.000235 |
| 62 | 0.993984 | 0.991802 | 0.99379 | 0.002181 | 0.000193 |
| 63 | 0.996681 | 0.995339 | 0.996533 | 0.001343 | 0.000148 |
| 64 | 0.998241 | 0.997445 | 0.998134 | 0.000796 | 0.000107 |
| 65 | 0.999105 | 0.99865 | 0.999032 | 0.000455 | 0.000073 |
| 66 | 0.999563 | 0.999313 | 0.999517 | 0.00025 | 0.000047 |
| 67 | 0.999796 | 0.999663 | 0.999767 | 0.000133 | 0.000028 |
| 68 | 0.999908 | 0.999841 | 0.999892 | 0.000068 | 0.000016 |
| 69 | 0.999961 | 0.999928 | 0.999952 | 0.000033 | 0.000009 |
| 70 | 0.999984 | 0.999968 | 0.999979 | 0.000016 | 0.000005 |
| 71 | 0.999994 | 0.999987 | 0.999991 | 0.000007 | 0.000002 |
| 72 | 0.999998 | 0.999995 | 0.999997 | 0.000003 | 0.000001 |
| 73 | 0.999999 | 0.999998 | 0.999999 | 0.000001 | 0 |
| 74 | 1.0 | 0.999999 | 1.0 | 0.000001 | 0 |
| 75 | 1.0 | 1.0 | 1.0 | 0 | 0 |

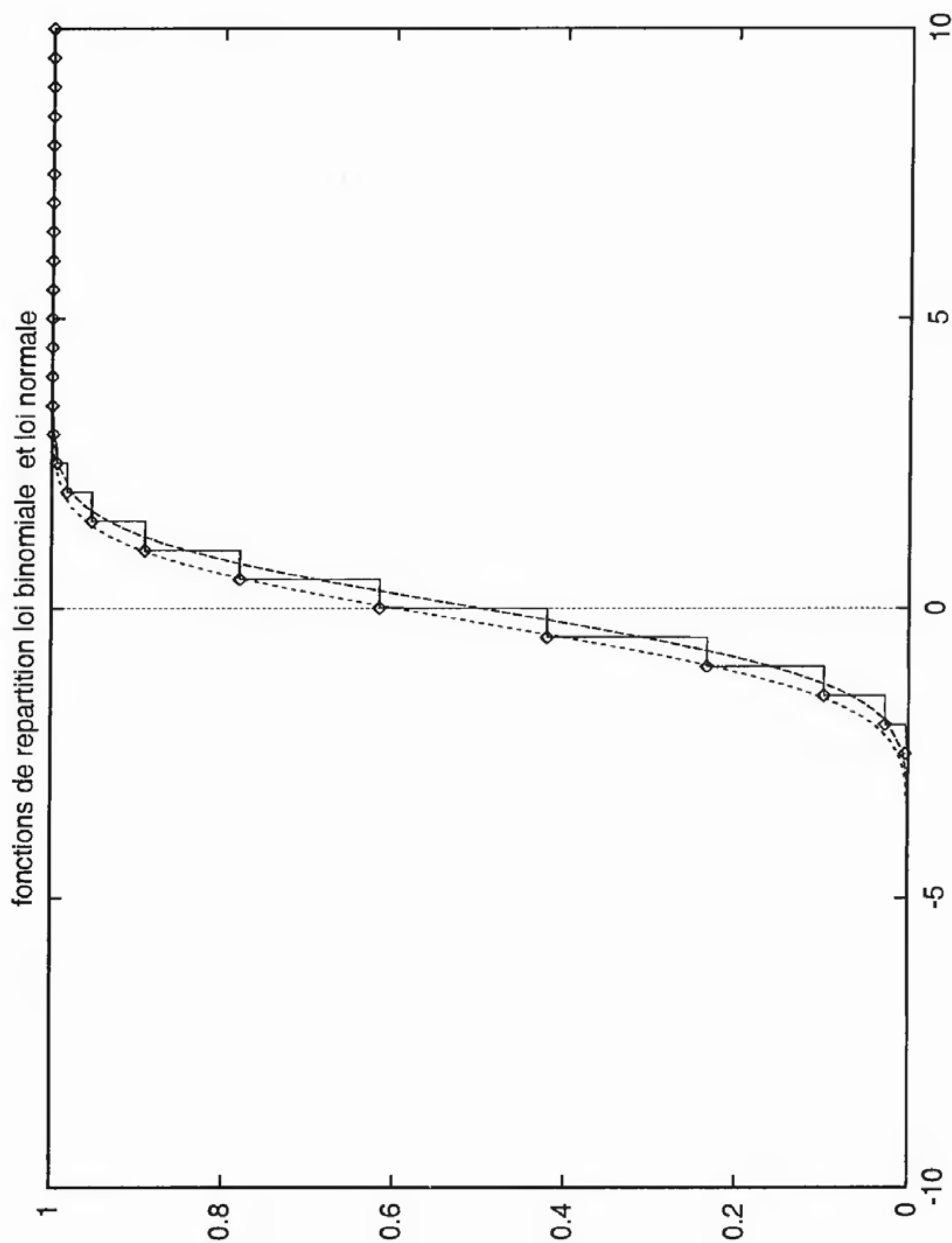



FIG. 8.2: $\mathcal{B}(25, 0.2)$ : $\widetilde{F}$ *(en escalier)*, $\Psi$ *(pointillé long)* , $\Psi^*$ *(pointillé court)*

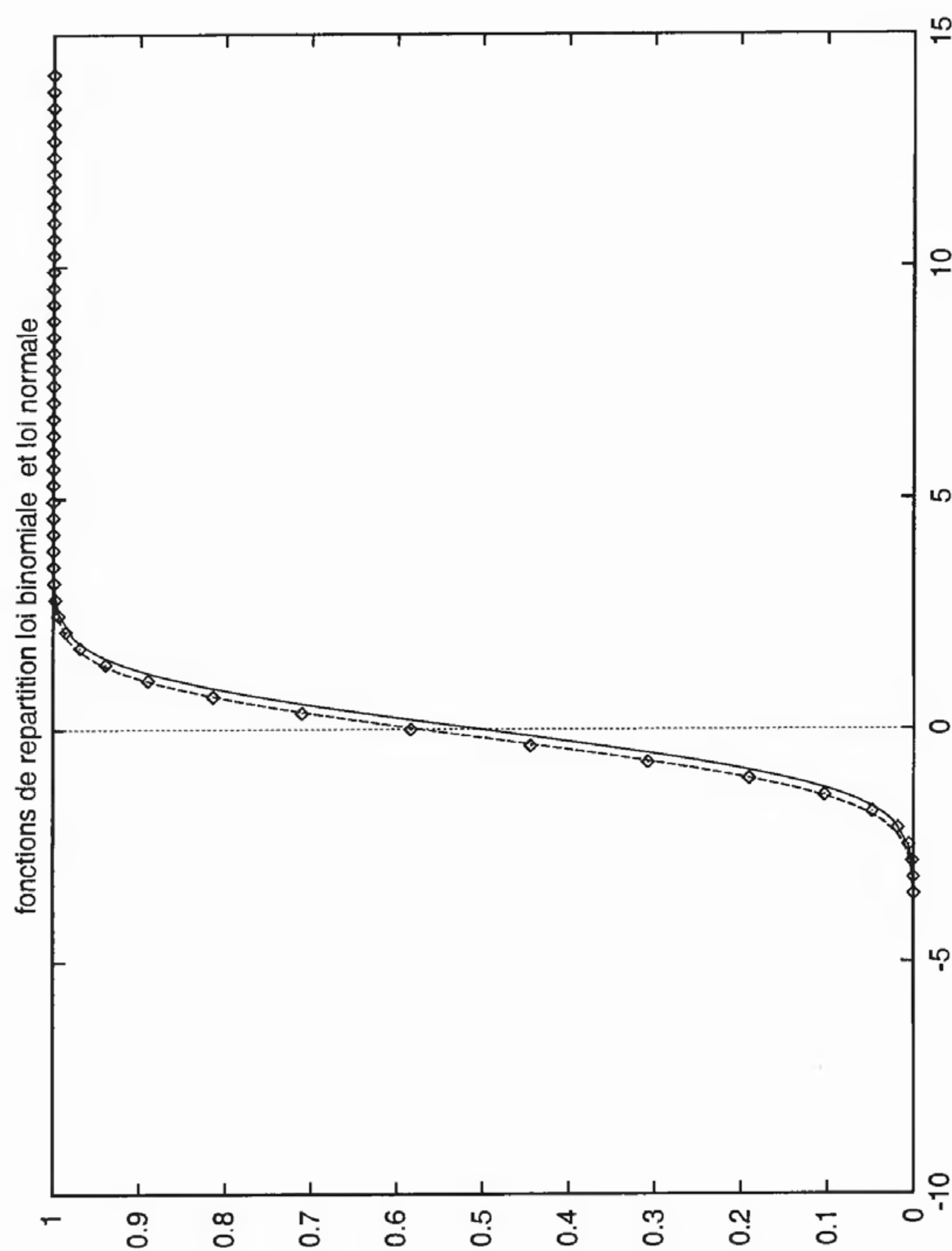



FIG. 8.3: $\mathcal{B}(50, 0.2)$ : *points* $(t_k, \widetilde{F}(t_k))$ $(0 \leqslant k \leqslant 50)$, $\Psi$, $\Psi^*$ *(pointillé)*

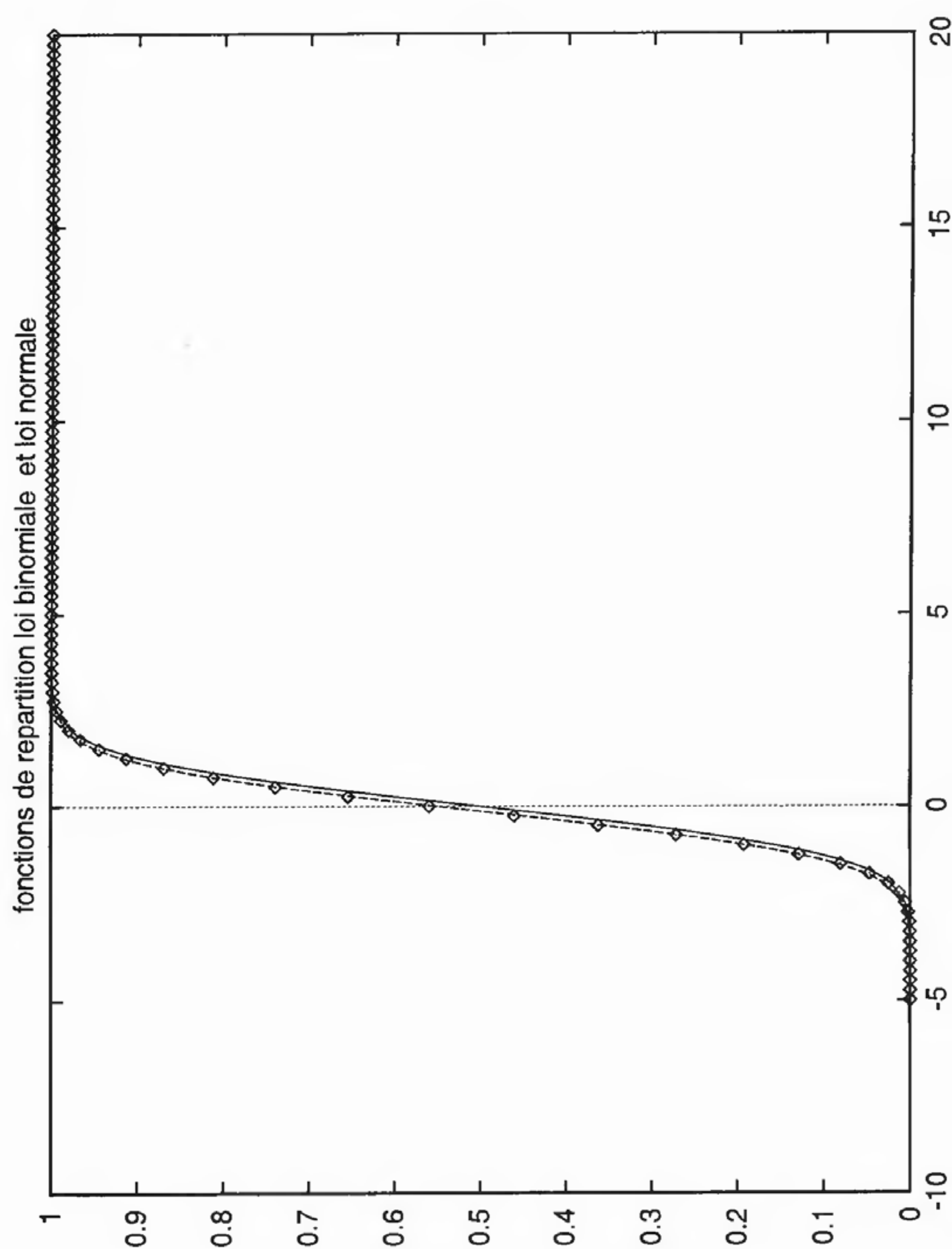



FIG. 8.4: $\mathcal{B}(100, 0.2)$ : *points* $(t_k, \widetilde{F}(t_k))$ $(0 \leqslant k \leqslant 100)$, $\Psi$, $\Psi^*$ *(pointillé)*

### 8.7.5 Application à la loi de Poisson.

**Théorème 8. 12.** *Pour tout $\lambda > 0$ soit $X_\lambda$ une v.a. qui suit une loi de Poisson $\mathcal{P}(\lambda)$, et $\widetilde{X_\lambda} = \frac{X_\lambda - \lambda}{\sqrt{\lambda}}$ la v.a. centrée réduite. Alors*

$$\lim_{\lambda \to +\infty} P(\widetilde{X_\lambda} \leqslant t) = \Psi(t) \quad \forall t \in \mathbb{R}\,. \tag{8.104}$$

*On notera encore simplement $\widetilde{X_\lambda} \Rightarrow \mathcal{N}(0,1)$ quand $\lambda \to +\infty$.*

**Démonstration.**

Supposons d'abord que $\lambda \to +\infty$ en étant de la forme $\lambda = n\gamma$, $n \in \mathbb{N}$, $n \to +\infty$, avec $\gamma > 0$ fixé. Pour tout $n$, il existe des v.a. indépendantes $Y_1, \cdots, Y_n$ suivant chacune la loi $\mathcal{P}(\gamma)$. Alors d'après l'exercice 8.14, $S_n = Y_1 + \cdots + Y_n$ suit la loi $\mathcal{P}(n\gamma)$. On peut donc supposer que $X_{n\gamma} = S_n$ pour tout $n$. On a alors directement d'après le Théorème central limite $\widetilde{X_{n\gamma}} = \widetilde{S_n} \Rightarrow \mathcal{N}(0,1)$ quand $n \to +\infty$. (8.104) est donc démontrée dans le cas où $\lambda = n\gamma$.

Notons que (8.104) s'écrit pour $t \in \mathbb{R}$ fixé

$$\lim_{\lambda \to +\infty} a_\lambda = \Psi(t) \tag{8.105}$$

où l'on a posé pour tout $x \geqslant 0$

$$a_x = e^{-x} \sum_{0 \leqslant k \leqslant x + t\sqrt{x}} \frac{x^k}{k!}\,. \tag{8.106}$$

Dans le cas général, soit donc $\varepsilon > 0$, et montrons que pour tout $t \in \mathbb{R}$ fixé, il existe $\lambda_1 > 0$ tel que

$$|a_\lambda - \Psi(t)| < \varepsilon \quad \forall\,\lambda \geqslant \lambda_1\,. \tag{8.107}$$

Soit $t \in \mathbb{R}$ fixé. Il existe $\alpha$ $(0 < \alpha < 1)$ tel que $\alpha(1 + \Psi(t)) + \alpha^2 < \varepsilon$, et $\gamma > 0$ tel que $e^\gamma < 1 + \alpha$ et $e^{-\gamma} > 1 - \alpha$, *i.e.*

$$0 < \gamma < \inf\left(\operatorname{Log}\,(1+\alpha)\,,\,|\operatorname{Log}\,(1-\alpha)|\right)\,.$$

Pour tout $\lambda > 0$, soit $n_\lambda$ la partie entière de $\frac{\lambda}{\gamma}$. On a

$$n_\lambda\,\gamma \leqslant \lambda < (n_\lambda + 1)\,\gamma. \tag{8.108}$$

Notons maintenant que la fonction $x \mapsto x + t\sqrt{x}$ est croissante et positive sur $[0, +\infty[$ si $t \geqslant 0$. Si $t < 0$, la dérivée de cette fonction s'annule pour $x = \frac{t^2}{4}$, et la fonction est croissante sur $[\frac{t^2}{4}, +\infty[$, positive sur $[t^2, +\infty[$. La fonction

$$x \mapsto e^x\,a_x = \sum_{0 \leqslant k \leqslant x + t\sqrt{x}} \frac{x^k}{k!} \tag{8.109}$$

est donc croissante sur $[0, +\infty[$ si $t \geqslant 0$ et sur $[t^2, +\infty[$ si $t < 0$. Dans les deux cas elle est croissante sur $[t^2, +\infty[$. Nous supposerons donc $\lambda \geqslant \lambda_0$ où $\lambda_0 = t^2 + \gamma$. Cette condition assure en effet que $\frac{\lambda}{\gamma} \geqslant \frac{t^2}{\gamma} + 1$, donc $n_\lambda \geqslant \frac{t^2}{\gamma}$

et $n_\lambda \gamma \geqslant t^2$. Pour tout $\lambda \geqslant \lambda_0$ la croissance de la fonction (8.109) sur $[t^2, +\infty[$ et l'inégalité (8.108) donnent donc

$$e^{n_\lambda \gamma} a_{n_\lambda \gamma} \leqslant e^\lambda a_\lambda \leqslant e^{(n_\lambda+1)\gamma} a_{(n_\lambda+1)\gamma}.$$

Comme $\lambda = n_\lambda \gamma + r$ $(0 \leqslant r < \gamma)$, il vient par simplification par $e^{n_\lambda \gamma}$

$$a_{n_\lambda \gamma} \leqslant e^r a_\lambda \leqslant e^\gamma a_{(n_\lambda+1)\gamma}$$

*i.e.*

$$e^{-r} a_{n_\lambda \gamma} \leqslant a_\lambda \leqslant e^{-r} e^\gamma a_{(n_\lambda+1)\gamma}.$$

Il vient enfin puisque $e^{-\gamma} < e^{-r} \leqslant 1$

$$e^{-\gamma} a_{n_\lambda \gamma} \leqslant a_\lambda \leqslant e^\gamma a_{(n_\lambda+1)\gamma}.$$

On a alors

$$(1-\alpha)\, a_{n_\lambda \gamma} < a_\lambda < (1+\alpha)\, a_{(n_\lambda+1)\gamma}. \tag{8.110}$$

Or d'après le premier cas traité,

$$\lim_{n\to+\infty} a_{n\gamma} = \Psi(t)\,.$$

Il existe donc $n_0$ tel que

$$\Psi(t) - \alpha < a_{n\gamma} < \Psi(t) + \alpha \quad \forall n \geqslant n_0\,.$$

Quand $\lambda \to +\infty$, $n_\lambda \to +\infty$ donc il existe $\lambda_1 \geqslant \lambda_0$ tel que $n_\lambda \geqslant n_0 \quad \forall \lambda \geqslant \lambda_1$. Alors $\Psi(t) - \alpha < a_{n_\lambda \gamma}$ et $a_{(n_\lambda+1)\gamma} < \Psi(t) + \alpha \quad \forall \lambda \geqslant \lambda_1$. On obtient donc d'après (8.110)

$$(1-\alpha)(\Psi(t)-\alpha) < a_\lambda < (1+\alpha)(\Psi(t)+\alpha) \quad \forall \lambda \geqslant \lambda_1\,.$$

D'où

$$-\alpha(1+\Psi(t)) + \alpha^2 < a_\lambda - \Psi(t) < \alpha(1+\Psi(t)) + \alpha^2 \quad \forall \lambda \geqslant \lambda_1\,.$$

(8.107) en résulte par définition de $\alpha$ car $-\alpha(1+\Psi(t)) - \alpha^2 < -\alpha(1+\Psi(t)) + \alpha^2$. □

Pour $\lambda \in \mathbb{R}$ grand, on peut donc approcher $\mathcal{P}(\lambda)$ par $\mathcal{N}(\lambda, \sqrt{\lambda})$. Une règle empirique est que l'approximation est acceptable si

$$\lambda \geqslant 6\,. \tag{8.111}$$

## 8.8 Appendice 1.

Nous donnons dans cet appendice une démonstration directe de l'approximation normale de la loi binomiale.

$S_n$ désigne une v.a. suivant la loi binomiale $\mathcal{B}(n,p)$ $(0 < p < 1)$, $\widetilde{S_n} = \frac{S_n - np}{\sqrt{npq}}$ est la v.a. centrée réduite, et $F_n$, $\widetilde{F_n}$ sont les fonctions de répartition de $S_n$, $\widetilde{S_n}$ respectivement.

**Lemme 8. 12.** *Les propriétés suivantes sont équivalentes.*
*(i) La suite* $(\widetilde{F_n})_{n\in\mathbb{N}}$ *converge simplement vers* $\Psi$ *sur* $\mathbb{R}$.
*(ii) La suite* $(\widetilde{F_n})_{n\in\mathbb{N}}$ *converge uniformément vers* $\Psi$ *sur* $\mathbb{R}$.
*(iii)* $\sup_{\alpha,\beta\in\mathbb{R},\alpha<\beta}\left|\widetilde{F_n}(\beta)-\widetilde{F_n}(\alpha)-(\Psi(\beta)-\Psi(\alpha))\right|\to 0$ *quand* $n\to+\infty$.
*(iv) Pour tous* $\alpha,\beta\in\mathbb{R}$, $\alpha<\beta$, *on a*

$$\left|\widetilde{F_n}(\beta)-\widetilde{F_n}(\alpha)-(\Psi(\beta)-\Psi(\alpha))\right|\to 0 \text{ quand } n\to+\infty\,.$$

**Démonstration.**
(i) $\Rightarrow$ (ii). Résulte directement du Théorème 8.9.
(ii) $\Rightarrow$ (iii). Soit $\varepsilon>0$. Il existe $n_0$ tel que $\sup_{x\in\mathbb{R}}|\widetilde{F_n}(x)-\Psi(x)|<\varepsilon\ \forall n\geqslant n_0$. Alors pour tous $\alpha,\beta\in\mathbb{R},\alpha<\beta$, et tout $n\geqslant n_0$

$$\left|\widetilde{F_n}(\beta)-\widetilde{F_n}(\alpha)-(\Psi(\beta)-\Psi(\alpha))\right|\leqslant\left|\widetilde{F_n}(\beta)-\Psi(\beta)\right|+\left|\widetilde{F_n}(\alpha)-\Psi(\alpha)\right|<2\varepsilon\,.$$

(iii) $\Rightarrow$ (iv). Trivial.
(iv) $\Rightarrow$ (i). Soit $\beta\in\mathbb{R}$ et $\varepsilon>0$. Comme

$$\lim_{T\to+\infty}\frac{1}{\sqrt{2\pi}}\int_{-T}^{T}e^{-\frac{t^2}{2}}\,dt=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}\,dt=1,$$

il existe $T>0$ tel que $-T<\beta$ et $\Psi(T)-\Psi(-T)\geqslant 1-\varepsilon$.
D'après (iv), il existe $n_0$ tel que

$$\left|\widetilde{F_n}(T)-\widetilde{F_n}(-T)-(\Psi(T)-\Psi(-T))\right|<\varepsilon\quad\forall n\geqslant n_0\,.$$

Mais alors $\widetilde{F_n}(T)-\widetilde{F_n}(-T)\geqslant 1-2\varepsilon\quad\forall n\geqslant n_0$ donc

$$\widetilde{F_n}(-T)\leqslant 2\varepsilon+\widetilde{F_n}(T)-1\leqslant 2\varepsilon\quad\forall n\geqslant n_0\,.$$

D'autre part, toujours d'après (iv), il existe $n_1$ tel que

$$\left|\widetilde{F_n}(\beta)-\widetilde{F_n}(-T)-(\Psi(\beta)-\Psi(-T))\right|<\varepsilon\quad\forall n\geqslant n_1\,.$$

On a alors pour $n\geqslant\sup(n_0,n_1)$ :

$$\left|\widetilde{F_n}(\beta)-\Psi(\beta)\right|\leqslant|\widetilde{F_n}(\beta)-F_n(-T)-(\Psi(\beta)-\Psi(-T))|$$
$$+\widetilde{F_n}(-T)+\Psi(-T)<4\varepsilon$$

puisque $\Psi(-T)\leqslant\Psi(T)-1+\varepsilon<\varepsilon$.
$\varepsilon>0$ étant arbitraire, $\lim_{n\to+\infty}\widetilde{F_n}(\beta)=\Psi(\beta)$. □

**Lemme 8. 13.** *Soient* $\alpha,\beta\in\mathbb{R}$, $\alpha<\beta$, *et*

$$E_n(\alpha,\beta)=\{k\in\{0,\cdots,n\};\ t_k\in]\alpha,\beta]\}\tag{8.112}$$

*où*

$$t_k=\frac{k-np}{\sqrt{npq}}\quad\forall k\in\{0,\cdots,n\}.\tag{8.113}$$

*Alors*

$$\sup_{k\in E_n(\alpha,\beta)}\left|\frac{C_n^k\,p^kq^{n-k}}{\frac{1}{\sqrt{2\pi npq}}\,e^{-\frac{t_k^2}{2}}}-1\right|\to 0 \text{ quand } n\to+\infty\,.\tag{8.114}$$

**Démonstration.**

• D'abord il existe $n_0$ tel que $k \neq 0, n \ \forall k \in E_n(\alpha,\beta)$, $\forall n \geqslant n_0$. En effet, pour $k = 0$ et $k = n$ (8.113) donne respectivement $t_0 = -\sqrt{\frac{np}{q}}$ et $t_n = \sqrt{\frac{nq}{p}}$. On a $t_0 \to -\infty$ et $t_n \to +\infty$ quand $n \to +\infty$, donc il existe $n_0$ tel que pour tout $n \geqslant n_0$ on ait $t_0 \notin ]\alpha,\beta]$ et $t_n \notin ]\alpha,\beta]$, *i.e.* $0, n \notin E_n(\alpha,\beta)$.

• La formule de Stirling s'écrit :

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n (1+\varepsilon_1(n)) \quad , \quad \lim_{n\to+\infty} \varepsilon_1(n) = 0 \ .$$

Elle donne pour $n \geqslant n_0$ et $k \in E_n(\alpha,\beta)$

$$C_n^k p^k q^{n-k} = \frac{\sqrt{2\pi n}\ \left(\frac{n}{e}\right)^n}{\sqrt{2\pi k}\ \left(\frac{k}{e}\right)^k\ \sqrt{2\pi\ell}\ \left(\frac{\ell}{e}\right)^\ell}\, p^k q^\ell\, (1+\varepsilon_2(n,k)) \qquad (8.115)$$

où $\ell = n - k$ et

$$1+\varepsilon_2(n,k) = \frac{1+\varepsilon_1(n)}{(1+\varepsilon_1(k))(1+\varepsilon_1(\ell))} \ .$$

• Maintenant

$$\sup_{k\in E_n(\alpha,\beta)} |\varepsilon_2(n,k)| \to 0 \text{ quand } n \to +\infty \ . \qquad (8.116)$$

En effet, la fonction $(x,y,z) \mapsto \frac{1+x}{(1+y)(1+z)}$ étant continue en $(0,0,0)$, pour tout $\varepsilon > 0$ il existe $\eta > 0$ tel que pour $|x|, |y|, |z| < \eta$, $\left|\frac{1+x}{(1+y)(1+z)} - 1\right| < \varepsilon$. Soit $n_1$ tel que $\varepsilon_1(n) < \eta \ \forall n \geqslant n_1$. Pour tout $k \in E_n(\alpha,\beta)$, on a

$$np + \alpha\sqrt{npq} < k \leqslant np + \beta\sqrt{npq} \quad \text{ et } \quad nq - \beta\sqrt{npq} \leqslant \ell < nq - \alpha\sqrt{npq}.$$

Comme $np + \alpha\sqrt{npq} \to +\infty$ et $nq - \beta\sqrt{npq} \to +\infty$ quand $n \to +\infty$, on en déduit qu'il existe $n_2$ tel que

$$k, \ell \geqslant n_1 \quad \forall k \in E_n(\alpha,\beta) \quad \forall n \geqslant n_2 \ .$$

On a alors

$$\sup_{k\in E_n(\alpha,\beta)} |\varepsilon_1(k)| < \eta \quad \text{ et } \quad \sup_{k\in E_n(\alpha,\beta)} |\varepsilon_1(\ell)| < \eta \quad \forall n \geqslant n_2$$

et donc

$$\sup_{k\in E_n(\alpha,\beta)} |\varepsilon_2(n,k)| < \varepsilon \qquad \forall n \geqslant \sup(n_0,n_2) \ .$$

• Ensuite considérons pour $n \geqslant n_0$ et $k \in E_n(\alpha,\beta)$

$$\pi(n,k) = \frac{\sqrt{2\pi n}\ \left(\frac{n}{e}\right)^n}{\sqrt{2\pi k}\ \left(\frac{k}{e}\right)^k\ \sqrt{2\pi\ell}\ \left(\frac{\ell}{e}\right)^\ell}\, p^k q^\ell = \frac{1}{\sqrt{2\pi npq}}\left(\frac{np}{k}\right)^{k+\frac{1}{2}}\left(\frac{nq}{\ell}\right)^{\ell+\frac{1}{2}} \ . \qquad (8.117)$$

(8.113) donne $\frac{k}{np} = 1 + t_k\sqrt{\frac{q}{np}}$ et $\frac{\ell}{nq} = 1 - t_k\sqrt{\frac{p}{nq}}$. On en déduit :

$$\begin{aligned}
\operatorname{Log}\left(\sqrt{2\pi npq}\,\pi(n,k)\right) &= \left(k+\frac{1}{2}\right)\operatorname{Log}\frac{np}{k} + \left(\ell+\frac{1}{2}\right)\operatorname{Log}\frac{nq}{\ell} \\
&= -\left(k+\frac{1}{2}\right)\operatorname{Log}\left(1+t_k\sqrt{\frac{q}{np}}\right) \\
&\qquad -\left(\ell+\frac{1}{2}\right)\operatorname{Log}\left(1-t_k\sqrt{\frac{p}{nq}}\right) \\
&= -\left(np+t_k\sqrt{npq}+\frac{1}{2}\right)\operatorname{Log}\left(1+t_k\sqrt{\frac{q}{np}}\right) \\
&\qquad -\left(nq-t_k\sqrt{npq}+\frac{1}{2}\right)\operatorname{Log}\left(1-t_k\sqrt{\frac{p}{nq}}\right).
\end{aligned}$$

Or

$$\sup_{k\in E_n(\alpha,\beta)} |t_k| \leqslant T = \sup(|\alpha|,|\beta|)\,. \tag{8.118}$$

Le développement limité $\operatorname{Log}(1+u) = u - \frac{u^2}{2} + \varphi(u)$ où $\varphi$ est une fonction qui est un $O(u^3)$ quand $u \to 0$, *i.e.* $\frac{\varphi(u)}{u^3}$ est bornée au voisinage de 0, donne

$$\begin{aligned}
&\operatorname{Log}\left(\sqrt{2\pi npq}\,\pi(n,k)\right) = \\
&\quad -\left(np+t_k\sqrt{npq}+\frac{1}{2}\right)\left(t_k\sqrt{\frac{q}{np}} - t_k^2\,\frac{q}{2np} + \varphi\left(t_k\sqrt{\frac{q}{np}}\right)\right) \\
&\quad -\left(nq-t_k\sqrt{npq}+\frac{1}{2}\right)\left(-t_k\sqrt{\frac{p}{nq}} - t_k^2\,\frac{p}{2nq} + \varphi\left(-t_k\sqrt{\frac{p}{nq}}\right)\right)
\end{aligned}$$

qui s'écrit

$$\operatorname{Log}\left(\sqrt{2\pi npq}\,\pi(n,k)\right) = -\frac{t_k^2}{2} + \psi(n,k) \tag{8.119}$$

avec

$$\begin{aligned}
\psi(n,k) = &-\frac{1}{2}t_k\left(\sqrt{\frac{q}{np}} - \sqrt{\frac{p}{nq}}\right) + \frac{1}{4}t_k^2\left(\frac{q}{np} - \frac{p}{nq}\right) + \frac{1}{2}t_k^3\left(q\sqrt{\frac{q}{np}} - p\sqrt{\frac{p}{nq}}\right) \\
&-\left(np+t_k\sqrt{npq}+\frac{1}{2}\right)\varphi\left(t_k\sqrt{\frac{q}{np}}\right) \\
&-\left(nq-t_k\sqrt{npq}+\frac{1}{2}\right)\varphi\left(-t_k\sqrt{\frac{p}{nq}}\right).
\end{aligned} \tag{8.120}$$

Comme $\varphi$ est un $O(u^3)$ quand $u \to 0$, il existe $c > 0$ et $M > 0$ tels que $|\varphi(u)| \leqslant M\,|u|^3 \quad \forall u\ ,\ 0 < |u| < c$ . Il existe $n_3$ tel que $\sup\left(T\sqrt{\frac{q}{np}}, T\sqrt{\frac{p}{nq}}\right) < c$ pour $n \geqslant n_3$. On a donc pour tout $n \geqslant \sup(n_0, n_3)$ d'après (8.118) :

$$\begin{aligned}
&\sup_{k\in E_n(\alpha,\beta)}\left|\varphi\left(t_k\sqrt{\frac{p}{nq}}\right)\right| \leqslant \frac{MT^3\left(\frac{p}{q}\right)^{\frac{3}{2}}}{n\sqrt{n}} \\
&\qquad\text{et}\ \sup_{k\in E_n(\alpha,\beta)}\left|\varphi\left(-t_k\sqrt{\frac{q}{np}}\right)\right| \leqslant \frac{MT^3\left(\frac{q}{p}\right)^{\frac{3}{2}}}{n\sqrt{n}}\,.
\end{aligned} \tag{8.121}$$

L'examen des divers termes de (8.120) compte tenu de (8.118) et (8.121) montre alors qu'il existe une constante $K$, dépendant seulement de $M, T, p, q$, telle que

$$\sup_{k \in E_n(\alpha,\beta)} |\psi(n,k)| \leqslant \frac{K}{\sqrt{n}} \quad \forall\, n \geqslant \sup(n_0, n_3). \tag{8.122}$$

Maintenant (8.119) s'écrit

$$\pi(n,k) = \frac{e^{-\frac{t_k^2}{2}}}{\sqrt{2\pi npq}}\, e^{\psi(n,k)} = \frac{e^{-\frac{t_k^2}{2}}}{\sqrt{2\pi npq}}\,(1 + \varepsilon_3(n,k)) \tag{8.123}$$

avec $\varepsilon_3(n,k) = e^{\psi(n,k)} - 1$ . Par continuité de l'application $x \mapsto e^x$ en $x = 0$, (8.122) prouve que

$$\sup_{k \in E_n(\alpha,\beta)} |\varepsilon_3(n,k)| \to 0 \text{ quand } n \to +\infty\,. \tag{8.124}$$

• D'après (8.115) et (8.123),

$$\begin{aligned} C_n^k p^k q^{n-k} = \pi(n,k)\,(1 + \varepsilon_2(n,k)) &= \frac{e^{-\frac{t_k^2}{2}}}{\sqrt{2\pi npq}}\,(1 + \varepsilon_2(n,k))(1 + \varepsilon_3(n,k)) \\ &= \frac{e^{-\frac{t_k^2}{2}}}{\sqrt{2\pi npq}}\,(1 + \varepsilon_4(n,k)) \end{aligned}$$

avec $\varepsilon_4(n,k) = \varepsilon_2(n,k) + \varepsilon_3(n,k) + \varepsilon_2(n,k)\,\varepsilon_3(n,k)$. Alors (8.114) résulte immédiatement de (8.116), (8.124). □

**Lemme 8. 14.** *Soit $f$ une fonction réelle continue sur l'intervalle fermé $[\alpha, \beta]$ $(\alpha, \beta \in \mathbb{R},\ \ \alpha < \beta)$. Pour tout $\nu \in \mathbb{N}^*$ soit $\tau_0^\nu = \alpha < \tau_1^\nu < \cdots < \tau_\nu^\nu = \beta$ une subdivision de $[\alpha, \beta]$ telle que $\sup_i |\tau_{i+1}^\nu - \tau_i^\nu| \to 0$ quand $\nu \to +\infty$. Enfin, pour tout $\nu$ et tout $i$ $(0 \leqslant i \leqslant \nu - 1)$, soit $\xi_i^\nu \in [\tau_i^\nu, \tau_{i+1}^\nu]$. Alors*

$$\int_\alpha^\beta f(t)\,dt = \lim_{\nu \to +\infty} \sum_{i=0}^{\nu-1} (\tau_{i+1}^\nu - \tau_i^\nu)\, f(\xi_i^\nu)\,. \tag{8.125}$$

**Démonstration.**

Soit $\varepsilon > 0$. Comme $f$ est continue sur l'intervalle compact $[\alpha, \beta]$, elle est uniformément continue, donc il existe $\eta > 0$ tel que $|f(t) - f(s)| < \varepsilon \quad \forall\, t, s\,, |t - s| < \eta$ . Or il existe $\nu_0$ tel que $\sup_i |\tau_{i+1}^\nu - \tau_i^\nu| < \eta \quad \forall\, \nu \geqslant \nu_0$ . On a alors pour tout $\nu \geqslant \nu_0$

$$\begin{aligned} \left| \sum_{i=0}^{\nu-1} (\tau_{i+1}^\nu - \tau_i^\nu)\, f(\xi_i^\nu) - \int_\alpha^\beta f(t)\,dt \right| &= \left| \sum_{i=0}^{\nu-1} \int_{\tau_i^\nu}^{\tau_{i+1}^\nu} (f(\xi_i^\nu) - f(t))\,dt \right| \\ &\leqslant \sum_{i=0}^{\nu-1} \int_{\tau_i^\nu}^{\tau_{i+1}^\nu} |f(\xi_i^\nu) - f(t)|\,dt \\ &< \varepsilon \sum_{i=0}^{\nu-1} (\tau_{i+1}^\nu - \tau_i^\nu) = \varepsilon(b - a). \end{aligned}$$

D'où le résultat puisque $\varepsilon > 0$ est arbitraire. □

**Lemme 8. 15.** *Soient $\alpha, \beta \in \mathbb{R}$, $\alpha < \beta$.*

$$P\left(\alpha < \widetilde{S_n} \leqslant \beta\right) = \widetilde{F_n}(\beta) - \widetilde{F_n}(\alpha) \to \Psi(\beta) - \Psi(\alpha) \quad \textit{quand } n \to +\infty\,. \tag{8.126}$$

**Démonstration.**

Soit $\varepsilon > 0$. D'après le Lemme 8.13 (formule (8.114)), il existe $n_0$ tel que pour tout $n \geqslant n_0$ et tout $k \in E_n(\alpha,\beta)$ ($E_n(\alpha,\beta)$ est défini par (8.112)),

$$\left| C_n^k p^k q^{n-k} - \frac{1}{\sqrt{2\pi npq}}\, e^{-\frac{t_k^2}{2}} \right| \leqslant \varepsilon \frac{1}{\sqrt{2\pi npq}}\, e^{-\frac{t_k^2}{2}} .$$

On a alors pour $n \geqslant n_0$ :

$$\left| P\left(\alpha < \widetilde{S_n} \leqslant \beta\right) - \sum_{k\in E_n(\alpha,\beta)} \frac{1}{\sqrt{2\pi npq}}\, e^{-\frac{t_k^2}{2}} \right| \leqslant \varepsilon \frac{1}{\sqrt{2\pi npq}} \sum_{k\in E_n(\alpha,\beta)} e^{-\frac{t_k^2}{2}} . \quad (8.127)$$

Or $E_n(\alpha,\beta)$ est un intervalle de $\mathbb{N}$ de la forme $[k_\alpha(n), k_\beta(n)]$.
Si $t_{k_\beta(n)} < \beta$, on introduit la subdivision de $[\alpha,\beta]$

$$\tau_0^\nu = \alpha < \tau_1^\nu = t_{k_\alpha(n)} < \tau_2^\nu = t_{k_\alpha(n)+1} < \cdots < \tau_{\nu-1}^\nu = t_{k_\beta(n)} < \tau_\nu^\nu = \beta$$

avec $k_\alpha(n) + \nu - 2 = k_\beta(n)$, *i.e.* $\nu = k_\beta(n) - k_\alpha(n) + 2$, et l'on pose

$$\xi_0^\nu = t_{k_\alpha(n)}, \xi_1^\nu = t_{k_\alpha(n)+1}, \cdots, \xi_{\nu-2}^\nu = t_{k_\beta(n)}, \xi_{\nu-1}^\nu = \beta.$$

Si $t_{k_\beta(n)} = \beta$, on introduit la subdivision de $[\alpha,\beta]$

$$\tau_0^\nu = \alpha < \tau_1^\nu = t_{k_\alpha(n)} < \tau_2^\nu = t_{k_\alpha(n)+1} < \cdots < \tau_\nu^\nu = t_{k_\beta(n)} = \beta$$

avec $k_\alpha(n) + \nu - 1 = k_\beta(n)$, *i.e.* $\nu = k_\beta(n) - k_\alpha(n) + 1$, et l'on pose

$$\xi_0^\nu = t_{k_\alpha(n)}, \xi_1^\nu = t_{k_\alpha(n)+1}, \cdots, \xi_{\nu-1}^\nu = t_{k_\beta(n)}.$$

Alors

$$\sum_{k\in E_n(\alpha,\beta)} \frac{1}{\sqrt{2\pi npq}}\, e^{-\frac{t_k^2}{2}} - \sum_{i=0}^{\nu-1} (\tau_{i+1}^\nu - \tau_i^\nu)\, e^{-\frac{(\xi_i^\nu)^2}{2}}$$

tend vers 0 quand $n \to +\infty$ puisque ces 2 sommes diffèrent pour tout $n$ d'au plus deux termes qui tendent vers 0. Or d'après le lemme 8.14

$$\lim_{n\to+\infty} \sum_{i=0}^{\nu-1} (\tau_{i+1}^\nu - \tau_i^\nu)\, e^{-\frac{(\xi_i^\nu)^2}{2}} = \frac{1}{\sqrt{2\pi}} \int_\alpha^\beta e^{-\frac{t^2}{2}}\, dt .$$

donc

$$\lim_{n\to+\infty} \sum_{k\in E_n(\alpha,\beta)} \frac{1}{\sqrt{2\pi npq}}\, e^{-\frac{t_k^2}{2}} = \frac{1}{\sqrt{2\pi}} \int_\alpha^\beta e^{-\frac{t^2}{2}}\, dt . \quad (8.128)$$

On déduit d'abord de (8.128) qu'il existe $n_1$ tel que pour tout $n \geqslant n_1$

$$\sum_{k\in E_n(\alpha,\beta)} \frac{1}{\sqrt{2\pi npq}}\, e^{-\frac{t_k^2}{2}} < \frac{2}{\sqrt{2\pi}} \int_\alpha^\beta e^{-\frac{t^2}{2}}\, dt < \frac{2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}}\, dt = 2 .$$

Par report dans (8.127), il vient

$$\left| P\left(\alpha < \widetilde{S_n} \leqslant \beta\right) - \sum_{k\in E_n(\alpha,\beta)} \frac{1}{\sqrt{2\pi npq}}\, e^{-\frac{t_k^2}{2}} \right| \leqslant 2\varepsilon \quad \forall\, n \geqslant \sup(n_0, n_1) .$$

Comme $\varepsilon > 0$ est arbitraire, cela donne

$$P\left(\alpha < \widetilde{S_n} \leqslant \beta\right) - \sum_{k \in E_n(\alpha,\beta)} \frac{1}{\sqrt{2\pi npq}} e^{-\frac{t_k^2}{2}} \to 0 \text{ quand } n \to +\infty . \qquad (8.129)$$

Il résulte alors de (8.128) que

$$\lim_{n\to+\infty} P\left(\alpha < \widetilde{S_n} \leqslant \beta\right) = \frac{1}{\sqrt{2\pi}} \int_\alpha^\beta e^{-\frac{t^2}{2}}\, dt \qquad (8.130)$$

qui est l'équation (8.126). □

**Théorème 8. 13 (de Moivre-Laplace).** *Soit $S_n$ une v.a. suivant la loi binomiale $\mathcal{B}(n,p)$ et $\widetilde{S_n} = \frac{S_n - np}{\sqrt{npq}}$ la v.a. centrée réduite. Alors*

$$\widetilde{S_n} \Rightarrow \mathcal{N}(0,1) \quad \text{quand } n \to +\infty.$$

**Démonstration.**
Résulte trivialement des lemmes 8.12 et 8.15. □

# 8.9 Appendice 2.

## 8.9.1 Démonstration du Théorème 8.3.

Par définition d'une partition, on a $\Lambda \neq \varnothing$.

Pour $\lambda \in \Lambda$, posons $s_\lambda = \sum_{i\in I_\lambda} u_i$. Montrons d'abord que $\sum_{i\in I} u_i \leqslant \sum_{\lambda\in\Lambda} s_\lambda$. Soit $F$ une partie finie de $I$ et $G = \{\lambda \in \Lambda\,; F \cap I_\lambda \neq \varnothing\}$. Il est clair que $G$ est fini. Alors

$$\sum_{i\in F} u_i = \sum_{\lambda\in G} \sum_{i \in F\cap I_\lambda} u_i \leqslant \sum_{\lambda\in G} s_\lambda \leqslant \sum_{\lambda\in\Lambda} s_\lambda.$$

Comme $F$ est une partie finie arbitraire de $I$, il vient $\sum_{i\in I} u_i \leqslant \sum_{\lambda\in\Lambda} s_\lambda$.

Montrons maintenant que $\sum_{\lambda\in\Lambda} s_\lambda \leqslant \sum_{i\in I} u_i$. S'il existe $\lambda \in \Lambda$ tel que $s_\lambda = +\infty$, alors pour tout $A > 0$, il existe une partie finie $G_A$ de $I_\lambda$ telle que $\sum_{i\in G_A} u_i > A$. Le nombre $A > 0$ étant arbitraire, on a alors $\sum_{i\in I} u_i = +\infty$, de sorte que l'inégalité $\sum_{\lambda\in\Lambda} s_\lambda \leqslant \sum_{i\in I} u_i$ est trivialement vérifié. On peut donc supposer que $s_\lambda < +\infty$ quel que soit $\lambda \in \Lambda$. Soit $G$ une partie finie $\neq \varnothing$ de $\Lambda$ et considérons $\sum_{\lambda\in G} s_\lambda$. Pour tout $\varepsilon > 0$ et tout $\lambda \in G$, il existe une partie finie $F_{\lambda,\varepsilon}$ de $I_\lambda$ telle que

$$s_\lambda - \frac{\varepsilon}{|G|} \leqslant \sum_{i\in F_{\lambda,\varepsilon}} u_i.$$

Soit $F_\varepsilon = \bigcup_{\lambda\in G} F_{\lambda,\varepsilon}$. C'est une partie finie de $I$, et l'on a

$$\sum_{i\in F_\varepsilon} u_i = \sum_{\lambda\in G} \sum_{i\in F_{\lambda,\varepsilon}} u_i \geqslant \sum_{\lambda\in G} s_\lambda - \varepsilon\ ,$$

donc

$$\sum_{i\in I} u_i \geqslant \sum_{\lambda\in G} s_\lambda - \varepsilon\ .$$

Comme $\varepsilon > 0$ est arbitraire, il vient

$$\sum_{i\in I} u_i \geqslant \sum_{\lambda\in G} s_\lambda\ .$$

$G$ étant une partie finie $\neq \varnothing$ arbitraire de $\Lambda$, on a donc $\sum_{\lambda\in\Lambda} s_\lambda \leqslant \sum_{i\in I} u_i$. □

### 8.9.2 Démonstration de la Proposition 8.2.

Supposons la famille sommable de somme $S$. Soit $\varepsilon > 0$ arbitraire, et $F_0$ une partie finie de $I$ telle que (8.54) soit vérifiée pour toute partie finie $F \subset I$ contenant $F_0$. Alors, pour toute partie finie $F \subset I$, on aura

$$\sum_{i \in F} u_i \leqslant \sum_{i \in F \cup F_0} u_i < S + \varepsilon$$

d'après (8.54) appliqué à $F \cup F_0$. On en déduit que

$$\sup_{\substack{F \subset I \\ F \text{fini}}} \sum_{i \in F} u_i \leqslant S + \varepsilon \,.$$

La somme (8.52) est donc finie. De plus

$$\sup_{\substack{F \subset I \\ F \text{fini}}} \sum_{i \in F} u_i \geqslant \sum_{i \in F_0} u_i > S - \varepsilon \,,$$

donc

$$\left| S - \sup_{\substack{F \subset I \\ F \text{fini}}} \sum_{i \in F} u_i \right| \leqslant \varepsilon \,.$$

Comme $\varepsilon$ est arbitraire,

$$S = \sup_{\substack{F \subset I \\ F \text{fini}}} \sum_{i \in F} u_i \,.$$

Réciproquement, supposons la somme (8.52) finie et désignons la par $S$. Par définition de $S$, pour toute partie finie $F \subset I$, on a

$$\sum_{i \in F} u_i \leqslant S \,.$$

Soit $\varepsilon > 0$. Par définition de $S$, il existe une partie finie $F_0 \subset I$ telle que

$$S - \varepsilon < \sum_{i \in F_0} u_i \leqslant S \,.$$

Pour toute partie finie $F \subset I$ contenant $F_0$, on aura alors

$$S - \varepsilon < \sum_{i \in F_0} u_i \leqslant \sum_{i \in F} u_i \leqslant S \,.$$

Ainsi $F_0$ est une partie finie de $I$ telle que (8.54) soit vérifiée pour toute partie finie $F \subset I$ contenant $F_0$. $\varepsilon > 0$ étant arbitraire, cela signifie que la famille est sommable de somme $S$. □

### 8.9.3 Démonstration du Théorème 8.4.

*Condition nécessaire.* Supposons la famille sommable de somme $S$. En prenant $\varepsilon = 1$ dans la définition 8.22, il existe une partie finie $F_0 \subset I$ telle que pour toute partie finie $F \subset I$ contenant $F_0$ on ait

$$\left| S - \sum_{i \in F} u_i \right| < 1 .$$

Soit maintenant $F$ une partie finie quelconque de $I$. On a

$$\left| S - \sum_{i \in F \cup F_0} u_i \right| < 1 . \tag{8.131}$$

D'autre part

$$\sum_{i \in F \cup F_0} u_i - \sum_{i \in F} u_i = \sum_{i \in F_0 \setminus F} u_i$$

donc

$$\left| \sum_{i \in F \cup F_0} u_i - \sum_{i \in F} u_i \right| \leqslant \sum_{i \in F_0 \setminus F} |u_i| \leqslant \sum_{i \in F_0} |u_i| . \tag{8.132}$$

On déduit de (8.131) et (8.132) que

$$\left| S - \sum_{i \in F} u_i \right| \leqslant 1 + \sum_{i \in F_0} |u_i| ,$$

et donc

$$\left| \sum_{i \in F} u_i \right| \leqslant A \tag{8.133}$$

avec

$$A = |S| + 1 + \sum_{i \in F_0} |u_i| .$$

$F$ étant une partie finie quelconque, l'ensemble des sommes finies est donc borné par $A$ :

$$\sup_{\substack{F \subset I \\ F \text{fini}}} \left| \sum_{i \in F} u_i \right| \leqslant A . \tag{8.134}$$

Nous allons montrer que cela implique

$$\sup_{\substack{F \subset I \\ F \text{fini}}} \sum_{i \in F} |u_i| \leqslant 2A , \tag{8.135}$$

et donc que la famille est absolument sommable.

Soit $F$ une partie finie quelconque de $I$. Introduisons

$$F_+ = \{ i \in F \,;\, u_i \geqslant 0 \} , \quad F_- = \{ i \in F \,;\, u_i < 0 \} .$$

Comme $F_+$ et $F_-$ sont des parties finies de $I$, on a d'après (8.134)

$$\left|\sum_{i\in F_+} u_i\right| \leqslant A, \quad \left|\sum_{i\in F_-} u_i\right| \leqslant A.$$

Or le signe de $u_i$ est fixe sur chacune des deux parties $F_+$ et $F_-$. Donc cela s'écrit encore

$$\sum_{i\in F_+} |u_i| \leqslant A, \quad \sum_{i\in F_-} |u_i| \leqslant A.$$

Par addition, on obtient (8.135) :

$$\sum_{i\in F} |u_i| \leqslant 2A.$$

*Condition suffisante.* Supposons la famille absolument sommable. Posons pour tout réel $x$

$$x^+ = \sup(x,0), \quad x^- = \sup(-x,0).$$

Alors

$$x = x^+ - x^-, \quad |x| = x^+ + x^-.$$

On a $0 \leqslant u_i^+ \leqslant |u_i| \; \forall i \in I$, Donc

$$\sup_{\substack{F\subset I\\ F\text{fini}}} \sum_{i\in F} u_i^+ \leqslant \sup_{\substack{F\subset I\\ F\text{fini}}} \sum_{i\in F} |u_i|.$$

Or

$$\sup_{\substack{F\subset I\\ F\text{fini}}} \sum_{i\in F} |u_i| < +\infty$$

puisque la famille $(u_i)_{i\in I}$ est absolument sommable. Donc

$$\sup_{\substack{F\subset I\\ F\text{fini}}} \sum_{i\in F} u_i^+ < +\infty.$$

La famille $(u_i^+)_{i\in I}$ est donc sommable. De même, on obtiendrait que la famille $(u_i^-)_{i\in I}$ est sommable. Or la famille $(u_i)_{i\in I}$ est la famille différence des deux familles $(u_i^+)_{i\in I}$ et $(u_i^-)_{i\in I}$. Elle est donc aussi sommable.

*Inégalité (8.56).* Soit $S = \sum_{i\in I} u_i$ la somme de la famille sommable. Soit $\varepsilon > 0$. Il existe une partie finie $F_0 \subset I$ telle que pour toute partie finie $F \subset I$ contenant $F_0$ on ait

$$\left|S - \sum_{i\in F} u_i\right| < \varepsilon.$$

Cela implique

$$\left|\,|S| - \left|\sum_{i\in F} u_i\right|\,\right| < \varepsilon$$

donc

$$|S| < \left|\sum_{i\in F} u_i\right| + \varepsilon \leqslant \sum_{i\in F} |u_i| + \varepsilon \leqslant \sum_{i\in I} |u_i| + \varepsilon\,.$$

$\varepsilon$ étant arbitraire,

$$|S| \leqslant \sum_{i\in I} |u_i|\,.$$

□

### 8.9.4 Démonstration du Théorème 8.5.

Pour tout $\lambda \in \Lambda$, la famille $(u_i)_{i\in I_\lambda}$ est une sous-famille de la famille sommable $(u_i)_{i\in I}$, donc elle est sommable. Soit $s_\lambda$ sa somme. On a

$$|s_\lambda| \leqslant \sum_{i\in I_\lambda} |u_i|$$

donc

$$\sum_{\lambda\in\Lambda} |s_\lambda| \leqslant \sum_{\lambda\in\Lambda}\sum_{i\in I_\lambda} |u_i| = \sum_{i\in I} |u_i| < +\infty$$

d'après le théorème d'associativité pour les familles sommables à termes $\geqslant 0$ (Th. 8.3). La famille $(s_\lambda)_{\lambda\in\Lambda}$ est donc sommable. Soit $T$ sa somme.

Nous allons maintenant montrer que $T$ est égal à la somme $S$ de la famille $(u_i)_{i\in I}$.

Soit $\varepsilon > 0$. Il existe une partie finie $F_0 \subset I$ telle que pour toute partie finie $F \subset I$ contenant $F_0$ on ait

$$\left|S - \sum_{i\in F} u_i\right| < \varepsilon\,. \tag{8.136}$$

Il existe de même une partie finie $G_0 \subset \Lambda$ telle que pour toute partie finie $G \subset \Lambda$ contenant $G_0$ on ait

$$\left|T - \sum_{\lambda\in G} s_\lambda\right| < \varepsilon\,. \tag{8.137}$$

On peut supposer $G_0 \neq \varnothing$.

Soit

$$G = G_0 \cup \{\lambda\in\Lambda\,;\, F_0\cap I_\lambda \neq \varnothing\}.$$

$G$ est une partie finie de $\Lambda$ et $G_0 \subset G$. Elle vérifie donc (8.137). De plus, notons que

$$F_0 = \bigcup_{\lambda\in G} (F_0\cap I_\lambda). \tag{8.138}$$

En effet, on a trivialement $\bigcup_{\lambda\in G}(F_0\cap I_\lambda) \subset F_0$. Mais

$$F_0 = F_0\cap I = F_0 \cap \bigcup_{\lambda\in\Lambda} I_\lambda = \bigcup_{\lambda\in\Lambda}(F_0\cap I_\lambda) = \bigcup_{\lambda\in\Lambda, F_0\cap I_\lambda\neq\varnothing} (F_0\cap I_\lambda) \subset \bigcup_{\lambda\in G}(F_0\cap I_\lambda).$$

Maintenant, pour tout $\lambda \in G$, il existe une partie finie $F_{\lambda,\varepsilon}$ de $I_\lambda$ telle que pour toute partie finie $H \subset I_\lambda$ contenant $F_{\lambda,\varepsilon}$ on ait

$$\left| s_\lambda - \sum_{i \in H} u_i \right| < \frac{\varepsilon}{|G|}. \tag{8.139}$$

Si l'on pose pour $\lambda \in G$

$$H_{\lambda,\varepsilon} = F_{\lambda,\varepsilon} \cup (F_0 \cap I_\lambda) \subset I_\lambda ,$$

$H_{\lambda,\varepsilon}$ est une partie finie de $I_\lambda$ qui contient $F_{\lambda,\varepsilon}$. Donc elle vérifie aussi (8.139). On en déduit que

$$\left| \sum_{\lambda \in G} s_\lambda - \sum_{\lambda \in G} \sum_{i \in H_{\lambda,\varepsilon}} u_i \right| < \varepsilon .$$

Alors

$$H_\varepsilon = \bigcup_{\lambda \in G} H_{\lambda,\varepsilon}$$

est une partie finie de $I$ et

$$\sum_{\lambda \in G} \sum_{i \in H_{\lambda,\varepsilon}} u_i = \sum_{i \in H_\varepsilon} u_i$$

donc

$$\left| \sum_{\lambda \in G} s_\lambda - \sum_{i \in H_\varepsilon} u_i \right| < \varepsilon . \tag{8.140}$$

Or $H_\varepsilon$ contient $\bigcup_{\lambda \in G} (F_0 \cap I_\lambda)$ qui d'après (8.138) est égal à $F_0$. Elle vérifie donc

$$\left| S - \sum_{i \in H_\varepsilon} u_i \right| < \varepsilon . \tag{8.141}$$

Par comparaison de (8.141) et (8.140), il vient

$$\left| S - \sum_{\lambda \in G} s_\lambda \right| < 2\varepsilon .$$

Mais $G$ vérifie (8.137), donc finalement

$$|S - T| < 3\varepsilon .$$

Comme $\varepsilon$ est arbitraire, on en déduit que $S = T$. □

## 8.10 Exercices.

**Exercice 8.1.**

Soit $\Omega$ un ensemble. Un ensemble de parties $\Lambda \subset \mathfrak{P}(\Omega)$ est appelé une *$\lambda$-classe* (la lettre $\lambda$ vient de l'anglais "lattice") si elle a les 3 propriétés suivantes :

(a) $\Omega \in \Lambda$;
(b) $A \setminus B \in \Lambda$ pour tous $A, B \in \Lambda$ tels que $B \subset A$;
(c) si $A_1, A_2, \cdots, A_n, \cdots \in \Lambda$ et $A_n \subset A_{n+1}$ $\forall n \in \mathbb{N}^*$, on a $\bigcup_{n=1}^{+\infty} A_n \in \Lambda$.

(i) Soit $\mathfrak{S} \subset \mathfrak{P}(\Omega)$ un ensemble de parties de $\Omega$. Montrer qu'il existe une plus petite $\lambda$-classe (pour l'inclusion) contenant $\mathfrak{S}$. On la notera $\Lambda(\mathfrak{S})$.

Dans toute la suite, on suppose que $\mathfrak{S}$ a la propriété suivante :

$$S_1 \cap S_2 \in \mathfrak{S} \quad \forall S_1, S_2 \in \mathfrak{S}. \tag{8.142}$$

(ii) Soit

$$\Lambda_1 = \{A \in \Lambda(\mathfrak{S})\ ;\ A \cap S \in \Lambda(\mathfrak{S}) \quad \forall S \in \mathfrak{S}\}.$$

Montrer que $\Lambda_1 = \Lambda(\mathfrak{S})$.

(iii) Soit

$$\Lambda_2 = \{A \in \Lambda(\mathfrak{S})\ ;\ A \cap T \in \Lambda(\mathfrak{S}) \quad \forall T \in \Lambda(\mathfrak{S})\}.$$

Montrer que $\Lambda_2 = \Lambda(\mathfrak{S})$.

(iv) Montrer que $\Lambda(\mathfrak{S})$ est la $\sigma$-algèbre engendrée par $\mathfrak{S}$.

**Indication.**

(i) Raisonner comme au lemme 8.2 et montrer que l'intersection $\Lambda = \bigcap_{i \in I} \Lambda_i$ d'une famille non vide quelconque $(\Lambda_i)_{i \in I}$ de $\lambda$-classes $\Lambda_i \subset \mathfrak{P}(\Omega)$ est une $\Lambda$-classe. En déduire le résultat en considérant la $\lambda$-classe intersection de la famille de toutes les $\lambda$-classes $\Lambda \subset \mathfrak{P}(\Omega)$ telles que $\mathfrak{S} \subset \Lambda$. Cette famille n'est pas vide puisqu'elle contient en particulier la $\Lambda$-classe $\mathfrak{P}(\Omega)$.

(ii) Vérifions que $\Lambda_1$ est une $\lambda$-classe.

- axiome (a) : $\Omega \in \Lambda_1$ car $\Omega \in \Lambda(\mathfrak{S})$ et $\Omega \cap S = S \in \Lambda(\mathfrak{S})\ \forall S \in \mathfrak{S}$.
- axiome (b) : Soient $A, B \in \Lambda_1$ avec $B \subset A$. Comme $A, B \in \Lambda(\mathfrak{S})$, on a aussi $A \setminus B \in \Lambda(\mathfrak{S})$ d'après l'axiome (b) valable pour $\Lambda(\mathfrak{S})$. Ensuite, pour tout $S \in \mathfrak{S}$, on a $(A \setminus B) \cap S = (A \cap S) \setminus (B \cap S) \in \Lambda(\mathfrak{S})$ puisque $A \cap S, B \cap S \in \Lambda(\mathfrak{S})$.
- axiome (c) : si $A_n \in \Lambda_1$ et $A_n \subset A_{n+1}$ $\forall n \in \mathbb{N}^*$, on a $\bigcup_{n=1}^{+\infty} A_n \in \Lambda(\mathfrak{S})$, et pour tout $S \in \mathfrak{S}$ , $\left(\bigcup_{n=1}^{+\infty} A_n\right) \cap S = \bigcup_{n=1}^{+\infty} (A_n \cap S) \in \Lambda(\mathfrak{S})$ d'après l'axiome (c) valable pour $\Lambda(S)$, puisque $A_n \cap S \in \Lambda(\mathfrak{S})$ et $A_n \cap S \subset A_{n+1} \cap S\ \forall n \in \mathbb{N}^*$. Donc $\bigcup_{n=1}^{+\infty} A_n \in \Lambda_1$.

Ainsi $\Lambda_1$ est une $\lambda$-classe. Or d'après (8.142), $\mathfrak{S} \subset \Lambda_1$. Donc $\Lambda(\mathfrak{S}) \subset \Lambda_1$ par définition de $\Lambda(\mathfrak{S})$. Mais $\Lambda_1 \subset \Lambda(\mathfrak{S})$ aussi par définition, d'où l'égalité.

(iii) Vérifions de même que $\Lambda_2$ est une $\lambda$-classe.

- axiome (a) : $\Omega \in \Lambda_2$ car $\Omega \in \Lambda(\mathfrak{S})$ et $\Omega \cap T = T \in \Lambda(\mathfrak{S})\ \forall T \in \Lambda(\mathfrak{S})$.
- axiome (b) : Soient $A, B \in \Lambda_2$ avec $B \subset A$. Comme $A, B \in \Lambda(\mathfrak{S})$, on a aussi $A \setminus B \in \Lambda(\mathfrak{S})$ d'après l'axiome (b) valable pour $\Lambda(\mathfrak{S})$. Ensuite, pour tout $T \in \Lambda(\mathfrak{S})$, on a $(A \setminus B) \cap T = (A \cap T) \setminus (B \cap T) \in \Lambda(\mathfrak{S})$ puisque $A \cap T, B \cap T \in \Lambda(\mathfrak{S})$.
- axiome (c) : si $A_n \in \Lambda_2$ et $A_n \subset A_{n+1}$ $\forall n \in \mathbb{N}^*$, on a on a $\bigcup_{n=1}^{+\infty} A_n \in \Lambda(\mathfrak{S})$, et pour tout $T \in \Lambda(\mathfrak{S})$ , $\left(\bigcup_{n=1}^{+\infty} A_n\right) \cap T = \bigcup_{n=1}^{+\infty} (A_n \cap T) \in \Lambda(\mathfrak{S})$ puisque $A_n \cap T \in \Lambda(\mathfrak{S})$ et $A_n \cap T \subset A_{n+1} \cap T\ \forall n \in \mathbb{N}^*$. Donc $\bigcup_{n=1}^{+\infty} A_n \in \Lambda_2$.

Ainsi $\Lambda_2$ est une $\lambda$-classe.

Soit maintenant $S \in \mathfrak{S}$. Par définition de $\Lambda_1$, $S \cap A \in \Lambda(\mathfrak{S})\ \forall A \in \Lambda_1$. Or d'après (ii) , $\Lambda_1 = \Lambda(\mathfrak{S})$. Donc $S \cap A \in \Lambda(\mathfrak{S})\ \forall A \in \Lambda(\mathfrak{S})$. Cela signifie que $S \in \Lambda_2$. $S \in \mathfrak{S}$ étant arbitraire, $\mathfrak{S} \subset \Lambda_2$. Mais $\Lambda_2 \subset \Lambda(\mathfrak{S})$ par définition, d'où l'égalité.

(iv) Soit $\sigma(\mathfrak{S})$ la $\sigma$-algèbre engendré par $\mathfrak{S}$. Toute $\sigma$-algèbre étant trivialement une $\lambda$-classe, on a $\Lambda(\mathfrak{S}) \subset \sigma(\mathfrak{S})$. Pour montrer que $\Lambda(\mathfrak{S}) = \sigma(\mathfrak{S})$, il suffit donc, par définition de $\sigma(\mathfrak{S})$, de démontrer que $\Lambda(\mathfrak{S})$ est une $\sigma$-algèbre contenant $\mathfrak{S}$.

- $\mathfrak{S} \subset \Lambda(\mathfrak{S})$ par définition de $\Lambda(\mathfrak{S})$.

• $\Omega \in \Lambda(\mathfrak{S})$ puisque $\Lambda(\mathfrak{S})$ est une $\lambda$-classe (axiome (a)).
• $\Lambda(\mathfrak{S})$ est stable par passage au complémentaire : si $A \in \Lambda(\mathfrak{S})$ on a en effet $A^c = \Omega \setminus A \in \Lambda(\mathfrak{S})$ d'après l'axiome (b) pour la $\lambda$-classe $\Lambda(\mathfrak{S})$.
• Montrons que $\Lambda(\mathfrak{S})$ est stable par union dénombrable. D'abord $\Lambda(\mathfrak{S})$ est stable par intersection finie : si $A, B \in \Lambda(\mathfrak{S})$, comme $\Lambda(\mathfrak{S}) = \Lambda_2$, on a immédiatement $A \cap B \in \Lambda(\mathfrak{S})$ par définition de $\Lambda_2$. Ensuite vérifions que $\Lambda(\mathfrak{S})$ est stable par union finie. Soient $A, B \in \Lambda(\mathfrak{S})$. On a $A \cup B = (A^c \cap B^c)^c$. Or comme $\Lambda(\mathfrak{S})$ est stable par passage au complémentaire, $A^c, B^c \in \Lambda(\mathfrak{S})$, donc $A^c \cap B^c \in \Lambda(\mathfrak{S})$ par stabilité par l'intersection, et $A \cup B \in \Lambda(\mathfrak{S})$ en utilisant à nouveau la stabilité par passage au complémentaire. Maintenant, soient $A_1, A_2, \cdots, A_n, \cdots \in \Lambda(\mathfrak{S})$ . On a $\bigcup_n A_n = \bigcup_n B_n$ avec $B_n = A_1 \cup \cdots \cup A_n$. Comme $\Lambda(\mathfrak{S})$ est stable par union finie, $B_n \in \Lambda(\mathfrak{S})\ \forall n \in \mathbb{N}^*$. D'après l'axiome (c) pour la $\lambda$-classe $\Lambda(\mathfrak{S})$, on a $\bigcup_n B_n \in \Lambda(\mathfrak{S})$, donc $\bigcup_n A_n = \bigcup_n B_n \in \Lambda(\mathfrak{S})$.
Ainsi $\Lambda(\mathfrak{S})$ est une $\sigma$-algèbre et par conséquent $\Lambda(\mathfrak{S}) = \sigma(\mathfrak{S})$.

**Exercice 8.2.**

Une colle est posée à 4 personnes. Elles doivent choisir une réponse parmi 3 réponses données. Comme les personnes n'ont aucune connaissance sur le sujet, elles répondent au hasard.

(i) Quelle est la probabilité $p_1$ que les 4 personnes donnent la même réponse?

(ii) Quelle est la probabilité $p_2$ que 2 réponses seulement apparaissent?

(iii) Quelle est la probabilité $p_3$ que les 3 réponses soient formulées?

**Indication.**

(i) Numérotons les personnes de 1 à 4 et les réponses possibles de 1 à 3. L'ensemble fondamental est alors l'ensemble de tous les quadruplets $\omega = (x_1, x_2, x_3, x_4)$ avec $x_i \in \{1,2,3\}\ \forall i$. On a $|\Omega| = 3^4 = 81$. Puisque les 4 personnes répondent au hasard, il y a équirépartition sur $\Omega$. $p_1 = \frac{3}{81} = \frac{1}{27} \approx 0.04$ .
(ii) Soit $A$ l'événement "3 personnes donnent une réponse, et la personne restante donne une autre réponse" et $B$ l'événement "2 personnes donnent une réponse, et les 2 personnes restantes donnent une autre réponse". L'événement "2 réponses seulement apparaissent" est $A \cup B$. On a $|A| = C_4^3 \cdot 3 \cdot 2 = 24$ (choix des 3 personnes qui donnent la même réponse × choix de leur réponse × choix de la réponse de la personne restante). Si l'on raisonne de même, pour $B$, on obtient : choix de 2 personnes qui donnent la même réponse × choix de leur réponse × choix de la réponse des 2 personnes restantes, mais dans ce cas il faut diviser par 2, car on compte 2 fois chaque $\omega \in B$ par ce procédé. D'où $|B| = (C_4^2 \cdot 3 \cdot 2)/2 = 18$. On pourrait aussi dire : choix de la personne qui donne la même réponse que la 1ère personne × choix de leur réponse × choix de la réponse des 2 personnes restantes. Cela donne $|B| = 3 \cdot 3 \cdot 2 = 18$. Maintenant, comme $A$ et $B$ sont incompatibles, $|A \cup B| = 42$. $p_2 = \frac{42}{81} = \frac{14}{27} \approx 0.52$ .
(iii) $p_3 = 1 - (p_1 + p_2) = \frac{12}{27} \approx 0.44$ .

**Exercice 8.3.**

On considère une assemblée de $n$ personnes ($n \leqslant 365$). Quelle est la probabilité $p$ qu'il y ait au moins deux personnes nées le même jour de l'année (on raisonne avec une année de 365 jours, et on suppose que le jour de naissance peut être un jour quelconque de l'année au hasard)? Que vaut $p$ pour $n = 23$? Pour $n = 50$?

**Indication.**

L'ensemble fondamental $\Omega$ est l'ensemble des $n$-uplets $(x_1, \cdots, x_n)$ avec $x_i \in \{1, \cdots, 365\}\ \forall i$, *i.e.* $\Omega = \{1, \cdots, 365\}^n$. Par hypothèse, il y a équiprobabilité. Si $A$ désigne l'événement "les $n$ personnes ont des jours de naissance deux-à-deux

distincts", on a donc

$$P(A) = \frac{365 \cdot 364 \cdots (365-(n-1))}{365^n} = \prod_{i=1}^{n} \frac{365-(i-1)}{365},$$

et $p = 1 - P(A)$. $p$ est une fonction croissante de $n$. Quelques valeurs :

| $n$ | 6 | 7 | 9 | 10 | 14 | 15 | 20 | 21 | 22 | 23 | 24 | 25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $p$ | 0.040 | 0.056 | 0.095 | 0.117 | 0.223 | 0.253 | 0.411 | 0.444 | 0.476 | 0.507 | 0.538 | 0.569 |

| $n$ | 30 | 34 | 35 | 40 | 41 | 47 | 49 | 50 | 53 | 57 | 60 | 70 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $p$ | 0.706 | 0.795 | 0.814 | 0.891 | 0.903 | 0.955 | 0.966 | 0.970 | 0.981 | 0.990 | 0.994 | 0.999 |

**Exercice 8.4.**

On utilise un jeu ordinaire de 52 cartes.

(i) On extrait 8 cartes au hasard. Quelle est la probabilité $p_1$ qu'il y ait les 4 as parmi les 8 cartes?

(ii) On commence par répartir le jeu en 4 tas correspondants aux 4 couleurs (trèfle, carreau, coeur et pique). On extrait ensuite 2 cartes au hasard dans chaque tas. Quelle est la probabilité $p_2$ qu'il y ait les 4 as parmi les 8 cartes obtenues?

**Indication.**

(i) L'ensemble fondamental $\Omega$ est l'ensemble des parties à 8 éléments extraites de l'ensemble des 52 cartes. $|\Omega| = C_{52}^8$. Il y a équiprobabilité. Le nombre de résultats favorables est ici le nombre de façons de compléter l'ensemble des 4 as en une partie à 8 éléments. C'est $C_{48}^4$. Donc $p_1 = \frac{C_{48}^4}{C_{52}^8} = \frac{2}{7735} \approx 0.0002$ .

(ii) L'ensemble fondamental $\Omega$ est maintenant l'ensemble des parties à 8 éléments extraites de l'ensemble des 52 cartes et formées de 2 cartes de chaque couleur. Pour chaque couleur il y a $C_{13}^2$ façons de choisir 2 cartes, donc $|\Omega| = (C_{13}^2)^4$. Il y a équiprobabilité. Un as d'une couleur étant donné, il y a 12 parties à 2 éléments extraites des 13 cartes de la couleur qui contiennent l'as. D'où

$$p_2 = \frac{(12)^4}{(C_{13}^2)^4} = \frac{16}{28561} \approx 0.0006 .$$

**Exercice 8.5.**

Une urne contient 10 boules dont 6 sont blanches et 4 sont rouges. On tire au hasard successivement 2 boules. Calculer, dans les deux cas avec remise et sans remise, les probabilités suivantes :

(i) Probabilité $p_1$ que les deux boules soient blanches.

(ii) Probabilité $p_2$ que les deux boules soient de la même couleur.

(iii) Probabilité $p_3$ que l'une au moins des deux boules soit blanche.

**Indication.**

(i) Soit $X$ la v.a. "nombre de boules blanches obtenues". On a $p_1 = P(X=2)$. S'il y a remise, $X$ suit $\mathcal{B}(n,p)$ avec $n=2$ et $p=\frac{6}{10}$ donc $P(X=2) = p^2 = \frac{36}{100}$. S'il n'y a pas remise, $X$ suit $\mathcal{H}(N, N_1, n)$ avec $N = 10, N_1 = 6, n = 2$, donc $P(X=2) = \frac{C_{N_1}^2 C_{N-N_1}^0}{C_N^2} = \frac{C_6^2}{C_{10}^2} = \frac{1}{3}$.

(ii) $p_2 = 1 - P(X=1)$. Avec remise, $P(X=1) = C_2^1 \frac{6}{10}\frac{4}{10} = \frac{48}{100}$ donc $p_2 = \frac{52}{100}$. Sans remise $P(X=1) = \frac{C_6^1 C_4^1}{C_{10}^2} = \frac{8}{15}$ donc $p_2 = \frac{7}{15} \approx 0.47$ .

(iii) $p_3 = 1 - P(X=0) = \frac{84}{100}$ avec remise et $p_3 = \frac{13}{15} \approx 0.87$ sans remise.

**Exercice 8.6.**

Trois machines $A, B, C$ produisent respectivement 70%, 20% et 10% des pièces fabriquées dans une usine. Le pourcentage de pièces défectueuses produites par chaque machine est respectivement 2%, 5% et 15% . On choisit une pièce au hasard dans la production, et on constate qu'elle est défectueuse. Quelle est la probabilité $p$ qu'elle vienne de la machine C?

**Indication.**

Soient $A$ (*resp.* $B, C$) l'événement "la pièce a été fabriquée par la machine A (*resp.* $B, C$)", et $D$ l'événement "la pièce est défectueuse". La formule de Bayes donne :

$$\begin{aligned} P(C|D) &= \frac{P(D|C)P(C)}{P(D|A)P(A) + P(D|B)P(B) + P(D|C)P(C)} \\ &= \frac{\frac{15}{100}\frac{10}{100}}{\frac{2}{100}\frac{70}{100} + \frac{5}{100}\frac{20}{100} + \frac{15}{100}\frac{10}{100}} = \frac{5}{13} \approx 0.38\ . \end{aligned}$$

**Exercice 8.7.**

On considère un serveur Internet. On modélise le nombre de connexions par jour par une v.a. $X$ sur un espace probabilisé $(\Omega, \mathfrak{A}, P)$. Des statistiques effectuées ont conduit à faire l'hypothèse que $X$ est de carré sommable, d'espérance 200 connexions par jour et d'écart-type 40. Donner un intervalle de confiance à 75 % pour $X$.

**Indication.**

Comme on ne connaît pas la loi de $X$, on utilise l'inégalité de Bienaymé-Tchebychev, sous la forme (8.83) : pour tout $t > 0$,

$$P\left(|X - E(X)| < t\,\sigma(X)\right) \quad \geqslant \quad 1 - \frac{1}{t^2}\ .$$

On fixe $t$ en sorte que $1 - \frac{1}{t^2} = 0.75$ , *i.e.* $t = 2$. Alors

$$P\left(|X - 200| < 2 \cdot 40\right) \quad \geqslant 0.75\,,$$

d'où l'intervalle de confiance $]120, 280[$.

**Exercice 8.8.**

Deux marques concurrentes $A$ et $B$ se partagent un marché. On constate que 25 % de la clientèle choisit $A$ et 75 % $B$.

(i) On effectue un sondage sur 100 personnes au hasard. Quelle est la loi suivie par la variable aléatoire $X$ = "nombre de personnes ayant acheté la marque $A$" ? Donner un intervalle de confiance à 95% pour $X$ (a) en utilisant la table de la loi binomiale, (b) en utilisant l'approximation normale, (c) en utilisant l'approximation normale avec correction de continuité.

(ii) Après une campagne publicitaire, on effectue une observation sur 100 personnes, et on constate que 31 ont choisi la marque $A$. La campagne publicitaire a-t-elle eu un impact ?

**Indication.**

(i) $X$ suit $\mathcal{B}(100, 0.25)$. (a) Si l'on note $F$ la fonction de répartition de $X$, on a d'après la table de la loi binomiale

$$P(17 \leqslant X \leqslant 33) = F(33) - F(16) \approx 0.9724 - 0.0211 = 0.9513\ .$$

L'intervalle $[17, 33]$, qui est centré sur l'espérance $E(X) = 25$, répond donc à la question.

(b) L'écart-type de $X$ est $\sqrt{100 \cdot \frac{1}{4} \cdot \frac{3}{4}} = \sqrt{\frac{300}{16}}$. Avec l'approximation normale, on approche la fonction de répartition $\widetilde{F}$ de $\widetilde{X} = \frac{X-25}{\sqrt{\frac{300}{16}}}$ par $\Psi$. On cherche donc $t > 0$ tel que $\Psi(t) - \Psi(-t) = 0.95$ . Cela s'écrit $\Phi(t) = 0.475$ , et donc d'après la table de la loi normale $t \approx 1.96$ . On a donc avec l'approximation normale

$$P(-1.96 < \widetilde{X} \leqslant 1.96) \approx \widetilde{F}(1.96) - \widetilde{F}(-1.96) \approx 0.95 .$$

Or

$$\begin{aligned} -1.96 < \widetilde{X} \leqslant 1.96 &\Leftrightarrow -1.96\sqrt{\frac{300}{16}} < X - 25 \leqslant 1.96\sqrt{\frac{300}{16}} \\ &\Leftrightarrow -8.4870 < X - 25 \leqslant 8.4870 \\ &\Leftrightarrow 16.5129 < X \leqslant 33.4870 \\ &\Leftrightarrow 17 \leqslant X \leqslant 33 . \end{aligned}$$

On retrouve donc $[17, 33]$.

(c) Avec l'approximation normale avec correction de continuité, on approche la fonction de répartition $\widetilde{F}$ de $\widetilde{X} = \frac{X-25}{\sqrt{\frac{300}{16}}}$ par $\Psi^*$, avec

$$\Psi^*(t) = \Psi(t + \frac{1}{2\sqrt{\frac{300}{16}}}) \approx \Psi(t + 0.1155).$$

On cherche donc $t > 0$ tel que $\Psi^*(t) - \Psi^*(-t) \approx 0.95$ . Cela s'écrit successivement (en arrondissant 0.1155 à 0.12) :

$$\Psi(t + 0.12) - \Psi(-t + 0.12) \approx 0.95$$

$$0.5 + \Phi(t + 0.12) - (0.5 - \Phi(t - 0.12)) \approx 0.95 \quad \text{(on a clairement } t > 0.12)$$

$$\Phi(t + 0.12) + \Phi(t - 0.12) \approx 0.95 . \tag{8.143}$$

La fonction $t \mapsto \Phi(t + 0.12) + \Phi(t - 0.12)$ est strictement croissante. D'après la table de la loi normale, elle vaut approximativement 0.9483 pour $t = 1.96$, 0.9495 pour $t = 1.97$, et 0.9507 pour $t = 1.98$. On peut donc prendre comme solution de (8.143) la valeur $t \approx 1.98$ . On a donc avec l'approximation normale avec correction de continuité

$$P(-1.98 < \widetilde{X} \leqslant 1.98) \approx 0.95 .$$

On obtient alors comme précédemment

$$\begin{aligned} -1.98 < \widetilde{X} \leqslant 1.98 &\Leftrightarrow -1.98\sqrt{\frac{300}{16}} < X - 25 \leqslant 1.98\sqrt{\frac{300}{16}} \\ &\Leftrightarrow -8.5736 < X - 25 \leqslant 8.5736 \\ &\Leftrightarrow 16.4263 < X \leqslant 33.5736 \\ &\Leftrightarrow 17 \leqslant X \leqslant 33 . \end{aligned}$$

On retrouve donc à nouveau $[17, 33]$.

(ii) Comme $31 \in [17, 33]$, on peut affirmer, avec un risque de 5% de se tromper, que la campagne n'a pas eu d'impact.

**Exercice 8.9.**

Deux candidats $A$ et $B$ uniquement se présentent à une élection présidentielle. Après la fermeture des bureaux de vote, les pourcentages de suffrages exprimés sont respectivement $p$ pour $A$ et $q = 1 - p$ pour $B$ (on néglige les bulletins blancs ou nuls). $p$ et $q$ sont inconnus avant dépouillement complet. Or un institut de sondage désire annoncer dès la fermeture des bureaux, sur la base d'un échantillon de $n$ bulletins pris au hasard dans l'ensemble des bureaux de vote, une *fourchette* du style *"$p = a$ à 3% près"*. Il faut entendre par là que, si $a$ désigne la moyenne observée dans l'échantillon,

$$P(|a - p| \leqslant 0.03) \geqslant 0.95 .$$

Quelle taille faut-il prévoir pour l'échantillon ? Même question avec 2%, avec 1%.

**Indication.**

Le nombre $X$ de suffrages exprimés pour $A$ dans un échantillon de taille $n$ suit la loi binomiale $\mathcal{B}(n, p)$. La moyenne observée dans l'échantillon est $a = \frac{X}{n}$. On veut donc que $P\left(\left|\frac{X}{n} - p\right| \leqslant 0.03\right) \geqslant 0.95$ ou encore en introduisant $\widetilde{X} = \frac{X-np}{\sqrt{npq}}$ :

$$P\left(\left|\widetilde{X}\right| \leqslant 0.03\sqrt{\frac{n}{pq}}\right) \geqslant 0.95 . \tag{8.144}$$

Avec l'approximation normale (on néglige la correction de continuité) on a

$$P\left(-t < \widetilde{X} \leqslant t\right) = 0.95$$

pour la valeur $t \approx 1.96$ . Pour réaliser (8.144), il suffit donc que

$$0.03\sqrt{\frac{n}{pq}} \geqslant 1.96 .$$

*i.e.*

$$\sqrt{n} \geqslant \frac{1.96}{0.03}\sqrt{pq} .$$

Or l'étude de la fonction $p \mapsto p(1-p)$ sur $[0, 1]$ montre que $pq \leqslant \frac{1}{4}$. Il suffit donc de réaliser

$$\sqrt{n} \geqslant \frac{1.96}{0.03}\frac{1}{2} = 32.67 .$$

La plus petite valeur qui convient est $n = 1068$.

Avec 2% et 1% on trouve respectivement $n = 2401$ et $n = 9604$.

**Exercice 8.10.**

Un examen consiste en un questionnaire à choix multiple. Il y a 50 questions, et pour chaque question le candidat doit choisir entre 4 réponses dont une seule est la bonne. Le score obtenu par un candidat est le nombre de questions où il a donné la bonne réponse. Les organisateurs de l'examen désirent introduire un score éliminatoire, tel que un candidat qui répondrait au hasard ait une chance sur 100 seulement d'atteindre ce score. Quel doit être ce score éliminatoire ? On effectuera les calculs : (i) avec la loi exacte, (ii) avec l'approximation normale, (iii) avec l'approximation normale avec correction de continuité.

**Indication.**

Le score $X$ d'un candidat qui répond au hasard à chaque question suit la loi binomiale $\mathcal{B}(50, 0.25)$. Le score éliminatoire $s_0$ est le plus petit entier $s$ tel que

$P(X \geqslant s) \leqslant 0.01$ . Or $P(X \geqslant s) = 1 - P(X \leqslant s-1)$. Donc $s_0$ est le plus petit entier $s$ tel que $P(X \leqslant s-1) \geqslant 0.99$ .
(i) La table 2 pour $n = 50$ et $p = 0.25$ montre que $s_0 - 1 = 20$. Donc $s_0 = 21$.
(ii) On a $X \leqslant s-1 \Leftrightarrow \widetilde{X} \leqslant t$ avec $\widetilde{X} = \frac{X-12.5}{\sqrt{\frac{150}{16}}}$ la variable centrée réduite et $t = \frac{s-1-12.5}{\sqrt{\frac{150}{16}}}$. En notant $\widetilde{F}$ la fonction de répartition de $\widetilde{X}$, l'équation

$$\widetilde{F}(t) = P(\widetilde{X} \leqslant t) \geqslant 0.99 \tag{8.145}$$

s'écrit avec l'approximation normale $\Psi(t) \geqslant 0.99$ . Le plus petit $t > 0$ tel que $\Psi(t) \geqslant 0.99$ est $\approx 2.33$ d'après la table 1. $s_0$ est donc le plus petit entier $s$ tel que $\frac{s-1-12.5}{\sqrt{\frac{150}{16}}} \geqslant 2.33$ , *i.e.* $s - 1 - 12.5 \geqslant 7.13$ , donc $s \geqslant 20.63$ . Cela donne $s_0 = 21$.
(iii) Avec l'approximation normale avec correction de continuité, l'équation (8.145) s'écrit $\Psi^*(t) = \Psi\left(t + \frac{1}{2\sqrt{\frac{150}{16}}}\right) \approx \Psi(t + 0.16) \geqslant 0.99$ . Le plus petit $t > 0$ tel que $\Psi^*(t) \geqslant 0.99$ est $\approx 2.33 - 0.16 = 2.17$ . $s_0$ est donc le plus petit entier $s$ tel que $\frac{s-1-12.5}{\sqrt{\frac{150}{16}}} \geqslant 2.17$ , *i.e.* $s - 1 - 12.5 \geqslant 6.64$ , donc $s \geqslant 20.14$ . D'où encore $s_0 = 21$.

**Exercice 8.11.**

La probabilité pour qu'il y ait une erreur typographique dans une page de livre est environ 3%. Quel serait le nombre maximum, au seuil de 5%, de pages à refaire dans un livre de 500 pages?

**Indication.**

Soit $X$ le nombre de pages présentant une erreur dans un livre de 500 pages. $X$ est une v.a. de loi $\mathcal{B}(500, 0.03)$. L'espérance de $X$ est 15 et son écart-type $\sqrt{15 \cdot 0.97} \approx 3.81$ . On cherche $t > 0$ tel que $P\left(-t < \widetilde{X} \leqslant t\right) = 0.95$ où $\widetilde{X}$ est la variable centrée réduite. On utilise l'approximation normale de la loi binomiale. Alors $t \approx 1.96$ , d'où le nombre maximum 22 pages puisque $15 + 3.81 \cdot 1.96 = 22.47$ .

**Exercice 8.12.**

Un étang contient un nombre inconnu $N$ de poissons. On effectue une première pêche de 185 poissons qu'on remet dans l'étang après les avoir marqués.

(i) Soit $X$ la variable aléatoire "nombre de poissons marqués obtenus lors d'une seconde pêche de 243 poissons". Quelle est la loi de $X$?

(ii) On effectue une seconde pêche de 243 poissons, et on constate que 41 sont marqués. Quelle est la valeur de $N$ qui rend maximum la probabilité $P(X = 41)$? On dit que cette valeur est *l'estimation par le maximum de vraisemblance.*

(iii) Quelle est l'estimation de $N$ par le maximum de vraisemblance dans le cas général d'une première pêche de $N_1$ poissons et d'une seconde pêche de $n$ poissons dont $k$ marqués? Peut-on prévoir "naïvement" l'estimation?

**Indication.**

(i) La loi de $X$ est la loi hypergéométrique $\mathcal{H}(N, 185, 243)$.
(ii) Soit $u_N = P(X = 41) = \frac{C_{185}^{41} C_{N-185}^{202}}{C_N^{243}}$. Considérons la fonction $N \mapsto u_N$. Alors

$$\frac{u_N}{u_{N-1}} = \frac{C_{N-185}^{202}}{C_{N-186}^{202}} \frac{C_{N-1}^{243}}{C_N^{243}} = \frac{N-185}{N-387} \frac{N-243}{N}$$

On a donc

$$\begin{aligned}\frac{u_N}{u_{N-1}} \geqslant 1 &\Leftrightarrow (N-185)(N-243) \geqslant N(N-387)\\ &\Leftrightarrow 41N \leqslant 185 \cdot 243\\ &\Leftrightarrow N \leqslant \left[\frac{185 \cdot 243}{41}\right] = 1096\end{aligned}$$

où le symbole $[x]$ désigne dans cet exercice la *partie entière* du nombre $x$. Donc $u_N$ est croissant pour $N \leqslant 1096$ et décroissant ensuite. Le maximum de vraisemblance est donc obtenu pour 1096.

(iii) On a $u_N = P(X=k) = \frac{C_{N_1}^k C_{N-N_1}^{n-k}}{C_N^n}$, donc

$$\frac{u_N}{u_{N-1}} = \frac{C_{N-N_1}^{n-k}}{C_{N-1-N_1}^{n-k}} \frac{C_{N-1}^n}{C_N^n} = \frac{N-N_1}{N-N_1-n+k} \frac{N-n}{N}$$

de sorte que

$$\begin{aligned}\frac{u_N}{u_{N-1}} \geqslant 1 &\Leftrightarrow (N-N_1)(N-n) \geqslant N(N-N_1-n+k)\\ &\Leftrightarrow -(N-N_1)n \geqslant N(-n+k)\\ &\Leftrightarrow N_1 n \geqslant Nk\\ &\Leftrightarrow N \leqslant \left[\frac{nN_1}{k}\right].\end{aligned}$$

Donc $u_N$ est croissant pour $N \leqslant \left[\frac{nN_1}{k}\right]$ et décroissant ensuite. Le maximum de vraisemblance est donc obtenu pour $\left[\frac{nN_1}{k}\right]$. On peut vérifier que c'est ce qu'on obtiendrait en disant simplement que les proportions sont conservées, puisque cela signifierait que

$$\frac{N_1}{N} = \frac{k}{n}.$$

**Exercice 8.13.**

Soit $X$ une v.a. qui suit la loi hypergéométrique $\mathcal{H}(N, N_1, n)$. Calculer
(i) l'espérance de $X$,
(ii) la variance de $X$.

**Indication.**

Avec $N_2 = N - N_1$, on a :

$$\begin{aligned}
E(X) &= \sum_{k\in\Omega_{(N,N_1,n)}} k\,\frac{C_{N_1}^{k}C_{N_2}^{n-k}}{C_N^n} \\
&= \frac{1}{C_N^n}\sum_{\substack{0\leqslant h_1\leqslant N_1\\ 0\leqslant h_2\leqslant N_2\\ h_1+h_2=n}} h_1\,C_{N_1}^{h_1}C_{N_2}^{h_2} \\
&\quad \text{(par définition de } \Omega_{(N,N_1,n)}\text{ , en posant } h_1 = k \text{ et } h_2 = n-k) \\
&= \frac{1}{C_N^n}\sum_{\substack{1\leqslant h_1\leqslant N_1\\ 0\leqslant h_2\leqslant N_2\\ h_1+h_2=n}} h_1\,C_{N_1}^{h_1}C_{N_2}^{h_2} \\
&= \frac{N_1}{C_N^n}\sum_{\substack{1\leqslant h_1\leqslant N_1\\ 0\leqslant h_2\leqslant N_2\\ h_1+h_2=n}} C_{N_1-1}^{h_1-1}C_{N_2}^{h_2} \quad (\text{car } h_1C_{N_1}^{h_1} = N_1C_{N_1-1}^{h_1-1}) \\
&= \frac{N_1}{C_N^n}\sum_{\substack{0\leqslant h_1'\leqslant N_1-1\\ 0\leqslant h_2\leqslant N_2\\ h_1'+h_2=n-1}} C_{N_1-1}^{h_1'}C_{N_2}^{h_2} \quad (\text{en posant } h_1' = h_1-1) \\
&= \frac{N_1}{C_N^n}\,C_{N-1}^{n-1} \quad (\text{formule } (8.45)) \\
&= n\,\frac{N_1}{N} \quad (\text{car } nC_N^n = NC_{N-1}^{n-1}) \\
&= np \quad (\text{en posant } p = \frac{N_1}{N}).
\end{aligned}$$

Dans le calcul ci-dessus, il est implicitement supposé que $N_1 \geqslant 1$. Le résultat obtenu $E(X) = np$ reste cependant valable aussi dans le cas $N_1 = 0$ puisque dans ce cas

$$\sum_{\substack{0\leqslant h_1\leqslant N_1\\ 0\leqslant h_2\leqslant N_2\\ h_1+h_2=n}} h_1\,C_{N_1}^{h_1}C_{N_2}^{h_2} = 0$$

et $p = \frac{N_1}{N} = 0$.

(ii)

$$
\begin{aligned}
\sum_{k\in\Omega_{(N,N_1,n)}} k(k-1)\,\frac{C_{N_1}^k C_{N_2}^{n-k}}{C_N^n} &= \frac{1}{C_N^n}\sum_{\substack{0\leqslant h_1\leqslant N_1\\ 0\leqslant h_2\leqslant N_2\\ h_1+h_2=n}} h_1(h_1-1)\,C_{N_1}^{h_1}C_{N_2}^{h_2}\\
&= \frac{1}{C_N^n}\sum_{\substack{2\leqslant h_1\leqslant N_1\\ 0\leqslant h_2\leqslant N_2\\ h_1+h_2=n}} h_1(h_1-1)\,C_{N_1}^{h_1}C_{N_2}^{h_2}\\
&= \frac{N_1(N_1-1)}{C_N^n}\sum_{\substack{2\leqslant h_1\leqslant N_1\\ 0\leqslant h_2\leqslant N_2\\ h_1+h_2=n}} C_{N_1-2}^{h_1-2}C_{N_2}^{h_2}\\
&\quad \text{(car } h_1(h_1-1)C_{N_1}^{h_1} = N_1(N_1-1)C_{N_1-2}^{h_1-2}\text{ )}\\
&= \frac{N_1(N_1-1)}{C_N^n}\sum_{\substack{0\leqslant h_1'\leqslant N_1-2\\ 0\leqslant h_2\leqslant N_2\\ h_1'+h_2=n-2}} C_{N_1-2}^{h_1'}C_{N_2}^{h_2}\\
&\quad \text{(en posant } h_1' = h_1-2\text{)}\\
&= \frac{N_1(N_1-1)}{C_N^n}\,C_{N-2}^{n-2}\\
&\quad \text{(formule (8.45))}\\
&= N_1(N_1-1)\frac{n(n-1)}{N(N-1)}\\
&\quad \text{(car } n(n-1)C_N^n = N(N-1)C_{N-2}^{n-2}\text{)}\\
&= np\,\frac{(N_1-1)(n-1)}{N-1}\quad \text{(en posant } p=\frac{N_1}{N}\text{)}.
\end{aligned}
$$

Dans le calcul ci-dessus, il est implicitement supposé que $N_1 \geqslant 2$. Le résultat obtenu reste cependant valable aussi dans le cas $N_1 = 0$ ou $N_1 = 1$ puisque dans ce cas

$$\sum_{\substack{0\leqslant h_1\leqslant N_1\\ 0\leqslant h_2\leqslant N_2\\ h_1+h_2=n}} h_1(h_1-1)\,C_{N_1}^{h_1}C_{N_2}^{h_2} = 0$$

et

$$np\,\frac{(N_1-1)(n-1)}{N-1} = 0.$$

Ainsi on a toujours

$$E(X^2) - E(X) = np\,\frac{(N_1-1)(n-1)}{N-1},$$

d'où

$$E(X^2) = np\,\frac{(N_1-1)(n-1)}{N-1} + E(X) = np\left(\frac{(N_1-1)(n-1)}{N-1} + 1\right).$$

Alors

$$\begin{aligned}
\operatorname{var}(X) &= E(X^2) - (E(X))^2 \\
&= E(X^2) - (np)^2 \\
&= np\left(\frac{(N_1-1)(n-1)}{N-1} + 1 - np\right) \\
&= np\left(\frac{(N_1-1)(n-1)}{N-1} + \frac{N-nN_1}{N}\right) \\
&= \frac{np}{N(N-1)}(N-N_1)(N-n) \\
&= npq\,\frac{N-n}{N-1}
\end{aligned}$$

en posant $q = \frac{N_2}{N} = 1 - p$. On notera que la formule pour $E(X)$ est la même que pour la loi binomiale $\mathcal{B}(n,p)$, mais la variance de $X$ est celle de la loi binomiale multipliée par le facteur $\frac{N-n}{N-1}$ qui tend vers 1 si $N \to +\infty$ avec $n$ fixé.

**Exercice 8.14.**

Soient $X_1, \cdots, X_n$ des v.a. indépendantes suivant des lois de Poisson $\mathcal{P}(\lambda_1), \cdots, \mathcal{P}(\lambda_n)$ respectivement. Montrer que la v.a. $X_1 + \cdots + X_n$ suit la loi de Poisson $\mathcal{P}(\lambda_1 + \cdots + \lambda_n)$.

**Indication.**

Pour $n = 2$, $X_1 + X_2$ est à valeurs dans $\mathbb{N}$ et pour tout $k \in \mathbb{N}$,

$$\begin{aligned}
P(X_1 + X_2 = k) &= \sum_{i=0}^{k} P(\{X_1 = i\} \cap \{X_2 = k-i\}) \\
&= \sum_{i=0}^{k} P(X_1 = i)\, P(X_2 = k-i) \\
&= \sum_{i=0}^{k} e^{-\lambda_1}\frac{\lambda_1^i}{i!}\, e^{-\lambda_2}\frac{\lambda_2^{k-i}}{(k-i)!} \\
&= e^{-(\lambda_1+\lambda_2)}\,\frac{1}{k!}\sum_{i=0}^{k} C_k^i \lambda_1^i \lambda_2^{k-i} \\
&= e^{-(\lambda_1+\lambda_2)}\,\frac{(\lambda_1+\lambda_2)^k}{k!}\,.
\end{aligned}$$

Donc $X_1 + X_2$ suit $\mathcal{P}(\lambda_1 + \lambda_2)$. Le cas général s'ensuit par récurrence sur $n$, puisque $X_n$ et $X_1 + \cdots + X_{n-1}$ sont indépendantes.

**Exercice 8.15.**

(Approximation de la loi binomiale par la loi de Poisson).

Soit $(X_n)_{n\geqslant 1}$ une suite de v.a. On suppose que pour chaque $n$, $X_n$ suit une loi binomiale $\mathcal{B}(n, p_n)$ $(0 < p_n < 1)$ et que $\lim_{n\to+\infty} np_n = \lambda > 0$. Montrer que pour tout $k \in \mathbb{N}$ fixé,

$$P(X_n = k) \to e^{-\lambda}\frac{\lambda^k}{k!} \quad \text{quand } n \to +\infty\,. \tag{8.146}$$

En déduire que

$$X_n \Rightarrow \mathcal{P}(\lambda) \quad \text{quand } n \to +\infty.$$

Pour $n$ "grand" et $p$ "petit", on a donc l'approximation $\mathcal{B}(n,p) \approx \mathcal{P}(\lambda)$ avec $\lambda = np$. Cela explique pourquoi la loi de Poisson apparaît dans de nombreux phénomènes naturels : elle intervient chaque fois qu'il s'agit de compter le nombre d'occurrences d'un grand nombre d'événements indépendants dont chacun isolément a une faible probabilité de réalisation.

Un résultat de Prokhorov ([18]) affirme en fait que :

$$\sum_{k=0}^{+\infty} \left| P(X_n = k) - e^{-\lambda}\frac{\lambda^k}{k!} \right| \leqslant \varrho_n(\lambda) \tag{8.147}$$

avec

$$\varrho_n(\lambda) = \frac{2\lambda}{n} \inf(2, \lambda). \tag{8.148}$$

**Indication.**

On a $p_n \sim \frac{\lambda}{n}$ quand $n \to +\infty$, donc

$$\begin{aligned} P(X_n = k) &= C_n^k p_n^k (1-p_n)^{n-k} \\ &\sim \frac{n(n-1)\cdots(n-k+1)}{k!}\left(\frac{\lambda}{n}\right)^k (1-p_n)^{n-k} \\ &\sim \frac{n(n-1)\cdots(n-k+1)}{n^k}\frac{\lambda^k}{k!}(1-p_n)^{n-k} \\ &\sim \frac{\lambda^k}{k!}(1-p_n)^{n-k} \end{aligned}$$

puisque $\frac{n(n-1)\cdots(n-k+1)}{n^k} \sim \frac{n^k}{n^k} = 1$ quand $n \to +\infty$. Maintenant

$$(1-p_n)^{n-k} = e^{y_n}$$

en posant

$$y_n = (n-k)\mathrm{Log}(1-p_n).$$

On a

$$y_n \sim n\left(\frac{-\lambda}{n}\right) = -\lambda.$$

Donc $e^{y_n} \to e^{-\lambda}$, et $e^{y_n} \sim e^{-\lambda}$. D'où (8.146). Maintenant, pour tout $t \in \mathbb{R}$,

$$P(X_n \leqslant t) = \sum_{0 \leqslant k \leqslant t} P(X_n = k) \to \sum_{0 \leqslant k \leqslant t} e^{-\lambda}\frac{\lambda^k}{k!} = \Pi(t) \quad \text{quand } n \to +\infty$$

où $\Pi$ désigne la fonction de répartition de la loi de Poisson $\mathcal{P}(\lambda)$. Donc $X_n \Rightarrow \mathcal{P}(\lambda) \quad$ quand $n \to +\infty$.

Les tableaux suivants donnent quelques exemples de comparaison des approximation normale (AN), normale avec correction de continuité (ANCC), et de Poisson pour la loi binomiale. Ces exemples illustrent le fait que l'approximation par la loi de Poisson est très bonne pour $np \leqslant 1$ et $n \geqslant 25$ et excellente pour pour $np \leqslant 1$ et $n \geqslant 100$.

Exemples de valeurs de la fonction de répartition $F$ de $\mathcal{B}(n,p)$ :
Approximation Normale (AN); Approximation Normale avec Correction de Continuité (ANCC); Approximation de Poisson.
$F(k), \Psi(\frac{k-np}{\sqrt{npq}}), \Psi^*(\frac{k-np}{\sqrt{npq}}) = \Psi(\frac{k+\frac{1}{2}-np}{\sqrt{npq}})$ et $\Pi(k) = \sum_{j=0}^{k} e^{-\lambda} \frac{\lambda^j}{j!}$ $(\lambda = np)$.

$n = 25$

$p = 0.2$ , $np = 5$ , $\gamma_{25}(0.2) = 0.272$ , $\rho_{25}(5) = 0.8$ .

| $k$ | $F$ | $\Psi$ | $\Psi^*$ | $\Pi$ | $F-\Psi$ | $F-\Psi^*$ | $F-\Pi$ |
|---|---|---|---|---|---|---|---|
| 0 | 0.003778 | 0.00621 | 0.012224 | 0.006738 | −0.002431 | −0.008446 | −0.00296 |
| 1 | 0.02739 | 0.02275 | 0.040059 | 0.040428 | 0.00464 | −0.012669 | −0.013037 |
| 2 | 0.098225 | 0.066807 | 0.10565 | 0.124652 | 0.031418 | −0.007424 | −0.026426 |
| 3 | 0.233993 | 0.158655 | 0.226627 | 0.265026 | 0.075338 | 0.007366 | −0.031032 |
| 4 | 0.420674 | 0.308538 | 0.401294 | 0.440493 | 0.112137 | 0.019381 | −0.019818 |
| 5 | 0.616689 | 0.5 | 0.598706 | 0.615961 | 0.116689 | 0.017983 | 0.000729 |
| 6 | 0.780035 | 0.691462 | 0.773373 | 0.762183 | 0.088573 | 0.006663 | 0.017852 |
| 7 | 0.890877 | 0.841345 | 0.89435 | 0.866628 | 0.049532 | −0.003473 | 0.024249 |
| 8 | 0.953226 | 0.933193 | 0.959941 | 0.931906 | 0.020033 | −0.006715 | 0.021319 |
| 9 | 0.982668 | 0.97725 | 0.987776 | 0.968172 | 0.005418 | −0.005107 | 0.014496 |
| 10 | 0.994445 | 0.99379 | 0.99702 | 0.986305 | 0.000655 | −0.002575 | 0.00814 |
| 11 | 0.99846 | 0.99865 | 0.999423 | 0.994547 | −0.00019 | −0.000963 | 0.003913 |
| 12 | 0.999631 | 0.999767 | 0.999912 | 0.997981 | −0.000136 | −0.00028 | 0.00165 |
| 13 | 0.999924 | 0.999968 | 0.999989 | 0.999302 | −0.000044 | −0.000065 | 0.000622 |
| 14 | 0.999986 | 0.999997 | 0.999999 | 0.999774 | −0.00001 | −0.000012 | 0.000213 |
| 15 | 0.999998 | 1.0 | 1.0 | 0.999931 | −0.000001 | −0.000001 | 0.000067 |
| 16 | 1.0 | 1.0 | 1.0 | 0.99998 | 0 | 0 | 0.00002 |
| 17 | 1.0 | 1.0 | 1.0 | 0.999995 | 0 | 0 | 0.000005 |
| 18 | 1.0 | 1.0 | 1.0 | 0.999999 | 0 | 0 | 0.000001 |

| | $\mathcal{B}(25\ ,\ 0.2)$ | AN | ANCC | $\mathcal{P}(5)$ |
|---|---|---|---|---|
| $P([3\ ,\ 7]) = F(7) - F(2)$ | 0.792652 | 0.774538 | 0.7887 | 0.741976 |
| $P([2\ ,\ 8]) = F(8) - F(1)$ | 0.925836 | 0.910443 | 0.919882 | 0.891479 |
| $P([1\ ,\ 9]) = F(9) - F(0)$ | 0.97889 | 0.97104 | 0.975551 | 0.961434 |
| $P([1\ ,\ 10]) = F(10) - F(0)$ | 0.990667 | 0.987581 | 0.984796 | 0.979567 |
| $P([0\ ,\ 11]) = F(11) - F(-1)$ | 0.99846 | 0.9973 | 0.996443 | 0.994547 |

$n = 25$

$p = 0.1$ , $np = 2.5$ , $\gamma_{25}(0.1) = 0.437333$ , $\rho_{25}(2.5) = 0.4$ .

| $k$ | $F$ | $\Psi$ | $\Psi^*$ | $\Pi$ | $F-\Psi$ | $F-\Psi^*$ | $F-\Pi$ |
|---|---|---|---|---|---|---|---|
| 0 | 0.07179 | 0.04779 | 0.091211 | 0.082085 | 0.023999 | −0.019421 | −0.010295 |
| 1 | 0.271206 | 0.158655 | 0.252493 | 0.287297 | 0.112551 | 0.018713 | −0.016091 |
| **2** | 0.537094 | 0.369441 | 0.5 | 0.543813 | 0.167653 | 0.037094 | −0.006719 |
| **3** | 0.763591 | 0.630559 | 0.747507 | 0.757576 | 0.133033 | 0.016084 | 0.006015 |
| 4 | 0.902006 | 0.841345 | 0.908789 | 0.891178 | 0.060662 | −0.006782 | 0.010828 |
| 5 | 0.9666 | 0.95221 | 0.97725 | 0.957979 | 0.01439 | −0.010649 | 0.008621 |
| 6 | 0.990524 | 0.990185 | 0.99617 | 0.985813 | 0.000339 | −0.005645 | 0.004711 |
| 7 | 0.997739 | 0.99865 | 0.999571 | 0.995753 | −0.000911 | −0.001832 | 0.001985 |
| 8 | 0.999542 | 0.999877 | 0.999968 | 0.99886 | −0.000334 | −0.000425 | 0.000683 |
| 9 | 0.999921 | 0.999993 | 0.999998 | 0.999723 | −0.000071 | −0.000077 | 0.000198 |
| 10 | 0.999988 | 1.0 | 1.0 | 0.999938 | −0.000011 | −0.000011 | 0.00005 |
| 11 | 0.999999 | 1.0 | 1.0 | 0.999987 | −0.000001 | −0.000001 | 0.000011 |
| 12 | 1.0 | 1.0 | 1.0 | 0.999998 | 0 | 0 | 0.000002 |

| | $\mathcal{B}(25\ ,\ 0.1)$ | AN | ANCC | $\mathcal{P}(2.5)$ |
|---|---|---|---|---|
| $P([2\ ,\ 3]) = F(3) - F(1)$ | 0.492385 | 0.471903 | 0.495015 | 0.470279 |
| $P([1\ ,\ 4]) = F(4) - F(0)$ | 0.830217 | 0.793554 | 0.817578 | 0.809093 |
| $P([0\ ,\ 5]) = F(5) - F(-1)$ | 0.9666 | 0.942394 | 0.9545 | 0.957979 |
| $P([0\ ,\ 6]) = F(6) - F(-1)$ | 0.990524 | 0.980369 | 0.973419 | 0.985813 |

$n = 25$

$p = 0.04$ , $np = 1$ , $\gamma_{25}(0.04) = 0.75379$ , $\rho_{25}(1) = 0.08$ .

| $k$ | $F$ | $\Psi$ | $\Psi^*$ | $\Pi$ | $F-\Psi$ | $F-\Psi^*$ | $F-\Pi$ |
|---|---|---|---|---|---|---|---|
| 0 | 0.360397 | 0.153717 | 0.304917 | 0.367879 | 0.20668 | 0.05548 | −0.007482 |
| **1** | 0.73581 | 0.5 | 0.695083 | 0.735759 | 0.23581 | 0.040727 | 0.000051 |
| 2 | 0.923517 | 0.846283 | 0.937107 | 0.919699 | 0.077234 | −0.01359 | 0.003818 |
| 3 | 0.983478 | 0.979387 | 0.994638 | 0.981012 | 0.004092 | −0.011159 | 0.002467 |
| 4 | 0.99722 | 0.9989 | 0.999823 | 0.99634 | −0.00168 | −0.002603 | 0.00088 |
| 5 | 0.999624 | 0.999978 | 0.999998 | 0.999406 | −0.000353 | −0.000373 | 0.000219 |
| 6 | 0.999958 | 1.0 | 1.0 | 0.999917 | −0.000041 | −0.000041 | 0.000042 |
| 7 | 0.999996 | 1.0 | 1.0 | 0.99999 | −0.000003 | −0.000003 | 0.000006 |

| | $\mathcal{B}(25\ ,\ 0.04)$ | AN | ANCC | $\mathcal{P}(1)$ |
|---|---|---|---|---|
| $P([0\ ,\ 1]) = F(1) - F(-1)$ | 0.73581 | 0.479387 | 0.63219 | 0.735759 |
| $P([0\ ,\ 2]) = F(2) - F(-1)$ | 0.923517 | 0.825669 | 0.874214 | 0.919699 |

$$n = 25$$

$$p = 0.01 \ , \quad np = 0.25 \ , \quad \gamma_{25}(0.01) = 1.57622 \ , \ \rho_{25}(0.25) = 0.005 \ .$$

| $k$ | $F$ | $\Psi$ | $\Psi^*$ | $\Pi$ | $F-\Psi$ | $F-\Psi^*$ | $F-\Pi$ |
|---|---|---|---|---|---|---|---|
| **0** | 0.777821 | 0.307651 | 0.692349 | 0.778801 | 0.47017 | 0.085473 | −0.000979 |
| **1** | 0.974241 | 0.934166 | 0.994008 | 0.973501 | 0.040075 | −0.019766 | 0.00074 |
| 2 | 0.998049 | 0.999782 | 0.999997 | 0.997839 | −0.001732 | −0.001947 | 0.000211 |
| 3 | 0.999893 | 1.0 | 1.0 | 0.999867 | −0.000106 | −0.000106 | 0.000026 |

| | $\mathcal{B}(25\ ,\ 0.01)$ | AN | ANCC | $\mathcal{P}(0.25)$ |
|---|---|---|---|---|
| $P([0\ ,\ 1]) = F(1) - F(-1)$ | 0.974241 | 0.928174 | 0.928174 | 0.973501 |
| $P([0\ ,\ 2]) = F(2) - F(-1)$ | 0.998049 | 0.99379 | 0.934163 | 0.997839 |

$$n = 50$$

$$p = 0.2 \ , \quad np = 10 \ , \quad \gamma_{50}(0.2) = 0.192333 \ , \ \rho_{50}(10) = 0.8 \ .$$

| $k$ | $F$ | $\Psi$ | $\Psi^*$ | $\Pi$ | $F-\Psi$ | $F-\Psi^*$ | $F-\Pi$ |
|---|---|---|---|---|---|---|---|
| 0 | 0.000014 | 0.000203 | 0.000391 | 0.000045 | −0.000189 | −0.000377 | −0.000031 |
| 1 | 0.000193 | 0.000731 | 0.001327 | 0.000499 | −0.000538 | −0.001134 | −0.000306 |
| 2 | 0.001285 | 0.002339 | 0.004005 | 0.002769 | −0.001053 | −0.002719 | −0.001483 |
| 3 | 0.005656 | 0.006664 | 0.010778 | 0.010336 | −0.001007 | −0.005121 | −0.004679 |
| 4 | 0.018496 | 0.016947 | 0.025915 | 0.029253 | 0.001549 | −0.007418 | −0.010756 |
| 5 | 0.048027 | 0.03855 | 0.055806 | 0.067086 | 0.009477 | −0.007778 | −0.019058 |
| 6 | 0.103398 | 0.07865 | 0.107962 | 0.130141 | 0.024749 | −0.004564 | −0.026743 |
| 7 | 0.19041 | 0.144422 | 0.18838 | 0.220221 | 0.045988 | 0.00203 | −0.02981 |
| 8 | 0.307332 | 0.23975 | 0.297942 | 0.33282 | 0.067582 | 0.00939 | −0.025488 |
| 9 | 0.44374 | 0.361837 | 0.429842 | 0.45793 | 0.081904 | 0.013899 | −0.014189 |
| **10** | 0.583559 | 0.5 | 0.570158 | 0.58304 | 0.083559 | 0.013401 | 0.00052 |
| 11 | 0.710668 | 0.638163 | 0.702058 | 0.696776 | 0.072504 | 0.008609 | 0.013891 |
| 12 | 0.813943 | 0.76025 | 0.81162 | 0.791556 | 0.053693 | 0.002323 | 0.022387 |
| 13 | 0.889413 | 0.855578 | 0.892038 | 0.864464 | 0.033836 | −0.002624 | 0.024949 |
| 14 | 0.939278 | 0.92135 | 0.944194 | 0.916542 | 0.017928 | −0.004916 | 0.022736 |
| 15 | 0.969197 | 0.96145 | 0.974085 | 0.95126 | 0.007747 | −0.004888 | 0.017937 |
| 16 | 0.985558 | 0.983053 | 0.989222 | 0.972958 | 0.002506 | −0.003663 | 0.0126 |
| 17 | 0.993739 | 0.993336 | 0.995995 | 0.985722 | 0.000403 | −0.002255 | 0.008017 |
| 18 | 0.997489 | 0.997661 | 0.998673 | 0.992813 | −0.000172 | −0.001184 | 0.004675 |
| 19 | 0.999068 | 0.999269 | 0.999609 | 0.996546 | −0.000201 | −0.00054 | 0.002522 |
| 20 | 0.999679 | 0.999797 | 0.999897 | 0.998412 | −0.000117 | −0.000217 | 0.001268 |
| 21 | 0.999898 | 0.99995 | 0.999976 | 0.9993 | −0.000051 | −0.000078 | 0.000597 |
| 22 | 0.99997 | 0.999989 | 0.999995 | 0.999704 | −0.000019 | −0.000025 | 0.000266 |
| 23 | 0.999992 | 0.999998 | 0.999999 | 0.99988 | −0.000006 | −0.000007 | 0.000112 |
| 24 | 0.999998 | 1.0 | 1.0 | 0.999953 | −0.000001 | −0.000001 | 0.000045 |
| 25 | 1.0 | 1.0 | 1.0 | 0.999982 | 0 | 0 | 0.000017 |

| | $\mathcal{B}(50\ ,\ 0.2)$ | AN | ANCC | $\mathcal{P}(10)$ |
|---|---|---|---|---|
| $P([8\ ,\ 12]) = F(12) - F(7)$ | 0.623533 | 0.615828 | 0.623241 | 0.571336 |
| $P([7\ ,\ 13]) = F(13) - F(6)$ | 0.786015 | 0.776928 | 0.784075 | 0.734323 |
| $P([6\ ,\ 14]) = F(14) - F(5)$ | 0.891251 | 0.8828 | 0.888388 | 0.849456 |
| $P([5\ ,\ 15]) = F(15) - F(4)$ | 0.950701 | 0.944503 | 0.94817 | 0.922007 |
| $P([4\ ,\ 16]) = F(16) - F(3)$ | 0.979902 | 0.976388 | 0.978444 | 0.962622 |
| $P([3\ ,\ 17]) = F(17) - F(2)$ | 0.992454 | 0.990997 | 0.99199 | 0.982953 |
| $P([2\ ,\ 18]) = F(18) - F(1)$ | 0.997296 | 0.99693 | 0.997346 | 0.992314 |

$$n = 50$$

$$p = 0.1\ ,\quad np = 5\ ,\quad \gamma_{50}(0.1) = 0.309241\ ,\ \rho_{50}(5) = 0.4\ .$$

| $k$ | $F$ | $\Psi$ | $\Psi^*$ | $\Pi$ | $F-\Psi$ | $F-\Psi^*$ | $F-\Pi$ |
|---|---|---|---|---|---|---|---|
| 0 | 0.005154 | 0.009211 | 0.016947 | 0.006738 | −0.004057 | −0.011793 | −0.001584 |
| 1 | 0.033786 | 0.029673 | 0.04948 | 0.040428 | 0.004113 | −0.015694 | −0.006641 |
| 2 | 0.111729 | 0.07865 | 0.119296 | 0.124652 | 0.033079 | −0.007567 | −0.012923 |
| 3 | 0.250294 | 0.172889 | 0.23975 | 0.265026 | 0.077405 | 0.010544 | −0.014732 |
| 4 | 0.431198 | 0.318676 | 0.406832 | 0.440493 | 0.112522 | 0.024367 | −0.009294 |
| **5** | 0.616123 | 0.5 | 0.593168 | 0.615961 | 0.116123 | 0.022955 | 0.000162 |
| 6 | 0.770227 | 0.681324 | 0.76025 | 0.762183 | 0.088903 | 0.009977 | 0.008043 |
| 7 | 0.877855 | 0.827111 | 0.880704 | 0.866628 | 0.050744 | −0.002848 | 0.011227 |
| 8 | 0.942133 | 0.92135 | 0.95052 | 0.931906 | 0.020782 | −0.008387 | 0.010226 |
| 9 | 0.975462 | 0.970327 | 0.983053 | 0.968172 | 0.005135 | −0.00759 | 0.00729 |
| 10 | 0.990645 | 0.990789 | 0.995239 | 0.986305 | −0.000143 | −0.004593 | 0.004341 |
| 11 | 0.99678 | 0.997661 | 0.998908 | 0.994547 | −0.000881 | −0.002128 | 0.002233 |
| 12 | 0.998995 | 0.999516 | 0.999797 | 0.997981 | −0.00052 | −0.000801 | 0.001014 |
| 13 | 0.999715 | 0.999919 | 0.999969 | 0.999302 | −0.000203 | −0.000254 | 0.000413 |
| 14 | 0.999926 | 0.999989 | 0.999996 | 0.999774 | −0.000062 | −0.00007 | 0.000152 |
| 15 | 0.999983 | 0.999999 | 1.0 | 0.999931 | −0.000016 | −0.000017 | 0.000052 |
| 16 | 0.999996 | 1.0 | 1.0 | 0.99998 | −0.000003 | −0.000003 | 0.000016 |
| 17 | 0.999999 | 1.0 | 1.0 | 0.999995 | 0 | 0 | 0.000005 |
| 18 | 1.0 | 1.0 | 1.0 | 0.999999 | 0 | 0 | 0.000001 |
| 19 | 1.0 | 1.0 | 1.0 | 1.0 | 0 | 0 | 0 |

| | $\mathcal{B}(50\ ,\ 0.1)$ | AN | ANCC | $\mathcal{P}(5)$ |
|---|---|---|---|---|
| $P([3\ ,\ 7]) = F(7) - F(2)$ | 0.766126 | 0.748461 | 0.761407 | 0.741976 |
| $P([2\ ,\ 8]) = F(8) - F(1)$ | 0.908347 | 0.891677 | 0.90104 | 0.891479 |
| $P([1\ ,\ 9]) = F(9) - F(0)$ | 0.970308 | 0.961116 | 0.966105 | 0.961434 |
| $P([0\ ,\ 10]) = F(10) - F(-1)$ | 0.990645 | 0.98845 | 0.990478 | 0.986305 |

$n = 50$

$p = 0.04$ , $np = 2$ , $\gamma_{50}(0.04) = 0.53301$ , $\rho_{50}(2.0) = 0.16$ .

| $k$ | $F$ | $\Psi$ | $\Psi^*$ | $\Pi$ | $F-\Psi$ | $F-\Psi^*$ | $F-\Pi$ |
|---|---|---|---|---|---|---|---|
| 0 | 0.129886 | 0.074457 | 0.139508 | 0.135335 | 0.055428 | −0.009622 | −0.005449 |
| 1 | 0.400481 | 0.235243 | 0.359108 | 0.406006 | 0.165238 | 0.041373 | −0.005524 |
| **2** | 0.676714 | 0.5 | 0.640892 | 0.676676 | 0.176714 | 0.035822 | 0.000038 |
| 3 | 0.860869 | 0.764757 | 0.860492 | 0.857123 | 0.096112 | 0.000377 | 0.003746 |
| 4 | 0.951029 | 0.925543 | 0.964402 | 0.947347 | 0.025486 | −0.013373 | 0.003682 |
| 5 | 0.98559 | 0.984809 | 0.99423 | 0.983436 | 0.000781 | −0.00864 | 0.002153 |
| 6 | 0.99639 | 0.998054 | 0.999418 | 0.995466 | −0.001663 | −0.003028 | 0.000924 |
| 7 | 0.999219 | 0.999846 | 0.999964 | 0.998903 | −0.000627 | −0.000745 | 0.000315 |
| 8 | 0.999852 | 0.999993 | 0.999999 | 0.999763 | −0.00014 | −0.000146 | 0.00009 |
| 9 | 0.999975 | 1.0 | 1.0 | 0.999954 | −0.000024 | −0.000024 | 0.000022 |
| 10 | 0.999996 | 1.0 | 1.0 | 0.999992 | −0.000003 | −0.000003 | 0.000005 |

| | $\mathcal{B}(50\ ,\ 0.04)$ | AN | ANCC | $\mathcal{P}(2.0)$ |
|---|---|---|---|---|
| $P([0\ ,\ 1]) = F(1) - F(-1)$ | 0.400481 | 0.220052 | 0.32351 | 0.406006 |
| $P([0\ ,\ 2]) = F(2) - F(-1)$ | 0.676714 | 0.484809 | 0.605293 | 0.676676 |
| $P([0\ ,\ 3]) = F(3) - F(-1)$ | 0.860869 | 0.749565 | 0.824893 | 0.857123 |

$n = 50$

$p = 0.02$ , $np = 1$ , $\gamma_{50}(0.02) = 0.776444$ , $\rho_{50}(1) = 0.04$ .

| $k$ | $F$ | $\Psi$ | $\Psi^*$ | $\Pi$ | $F-\Psi$ | $F-\Psi^*$ | $F-\Pi$ |
|---|---|---|---|---|---|---|---|
| 0 | 0.36417 | 0.156211 | 0.306753 | 0.367879 | 0.207959 | 0.057417 | −0.003709 |
| **1** | 0.735771 | 0.5 | 0.693247 | 0.735759 | 0.235771 | 0.042524 | 0.000013 |
| 2 | 0.921572 | 0.843789 | 0.935143 | 0.919699 | 0.077783 | −0.01357 | 0.001874 |
| 3 | 0.982242 | 0.978324 | 0.994221 | 0.981012 | 0.003918 | −0.011979 | 0.00123 |
| 4 | 0.99679 | 0.998779 | 0.999797 | 0.99634 | −0.001988 | −0.003006 | 0.00045 |
| 5 | 0.999522 | 0.999973 | 0.999997 | 0.999406 | −0.000451 | −0.000475 | 0.000116 |
| 6 | 0.99994 | 1.0 | 1.0 | 0.999917 | −0.000059 | −0.00006 | 0.000023 |
| 7 | 0.999994 | 1.0 | 1.0 | 0.99999 | −0.000006 | −0.000006 | 0.000004 |

| | $\mathcal{B}(50\ ,\ 0.02)$ | AN | ANCC | $\mathcal{P}(1)$ |
|---|---|---|---|---|
| $P([0\ ,\ 1]) = F(1) - F(-1)$ | 0.735771 | 0.478324 | 0.62839 | 0.735759 |
| $P([0\ ,\ 2]) = F(2) - F(-1)$ | 0.921572 | 0.822113 | 0.870286 | 0.919699 |
| $P([0\ ,\ 3]) = F(3) - F(-1)$ | 0.982242 | 0.956648 | 0.929364 | 0.981012 |

$n = 50$

$p = 0.01$ , $np = 0.5$ , $\gamma_{50}(0.01) = 1.11456$ , $\rho_{50}(0.5) = 0.01$ .

| $k$ | $F$ | $\Psi$ | $\Psi^*$ | $\Pi$ | $F-\Psi$ | $F-\Psi^*$ | $F-\Pi$ |
|---|---|---|---|---|---|---|---|
| **0** | 0.605006 | 0.238645 | 0.5 | 0.606531 | 0.366361 | 0.105006 | −0.001524 |
| **1** | 0.910565 | 0.761355 | 0.922391 | 0.909796 | 0.149209 | −0.011826 | 0.000769 |
| 2 | 0.986183 | 0.983497 | 0.997763 | 0.985612 | 0.002686 | −0.01158 | 0.00057 |
| 3 | 0.998404 | 0.99981 | 0.99999 | 0.998248 | −0.001405 | −0.001586 | 0.000155 |
| 4 | 0.999854 | 1.0 | 1.0 | 0.999828 | −0.000145 | −0.000145 | 0.000026 |

| | $\mathcal{B}(50\ ,\ 0.01)$ | AN | ANCC | $\mathcal{P}(0.5)$ |
|---|---|---|---|---|
| $P([0\ ,\ 1]) = F(1) - F(-1)$ | 0.910565 | 0.744852 | 0.844782 | 0.909796 |
| $P([0\ ,\ 2]) = F(2) - F(-1)$ | 0.986183 | 0.966994 | 0.920154 | 0.985612 |
| $P([0\ ,\ 3]) = F(3) - F(-1)$ | 0.998404 | 0.983307 | 0.922381 | 0.998248 |

$$n = 100$$

$$p = 0.1\ \ ,\ \ np = 10\ \ ,\ \ \gamma_{100}(0.1) = 0.218667\ ,\ \rho_{100}(10) = 0.4\ .$$

| $k$ | $F$ | $\Psi$ | $\Psi^*$ | $\Pi$ | $F-\Psi$ | $F-\Psi^*$ | $F-\Pi$ |
|---|---|---|---|---|---|---|---|
| 0 | 0.000027 | 0.000429 | 0.000771 | 0.000045 | −0.000402 | −0.000744 | −0.000018 |
| 1 | 0.000322 | 0.00135 | 0.002303 | 0.000499 | −0.001028 | −0.001981 | −0.000177 |
| 2 | 0.001945 | 0.00383 | 0.00621 | 0.002769 | −0.001885 | −0.004264 | −0.000824 |
| 3 | 0.007836 | 0.009815 | 0.01513 | 0.010336 | −0.001978 | −0.007293 | −0.002499 |
| 4 | 0.023711 | 0.02275 | 0.033377 | 0.029253 | 0.000961 | −0.009665 | −0.005541 |
| 5 | 0.057577 | 0.04779 | 0.066807 | 0.067086 | 0.009787 | −0.00923 | −0.009509 |
| 6 | 0.117156 | 0.091211 | 0.121673 | 0.130141 | 0.025944 | −0.004516 | −0.012985 |
| 7 | 0.206051 | 0.158655 | 0.202328 | 0.220221 | 0.047396 | 0.003722 | −0.014169 |
| 8 | 0.320874 | 0.252493 | 0.308538 | 0.33282 | 0.068381 | 0.012336 | −0.011945 |
| 9 | 0.45129 | 0.369441 | 0.433816 | 0.45793 | 0.081849 | 0.017474 | −0.006639 |
| **10** | 0.583156 | 0.5 | 0.566184 | 0.58304 | 0.083156 | 0.016972 | 0.000116 |
| 11 | 0.703033 | 0.630559 | 0.691462 | 0.696776 | 0.072474 | 0.011571 | 0.006257 |
| 12 | 0.801821 | 0.747507 | 0.797672 | 0.791556 | 0.054314 | 0.004149 | 0.010265 |
| 13 | 0.876123 | 0.841345 | 0.878327 | 0.864464 | 0.034778 | −0.002204 | 0.011659 |
| 14 | 0.927427 | 0.908789 | 0.933193 | 0.916542 | 0.018638 | −0.005765 | 0.010886 |
| 15 | 0.960109 | 0.95221 | 0.966623 | 0.95126 | 0.0079 | −0.006514 | 0.00885 |
| 16 | 0.979401 | 0.97725 | 0.98487 | 0.972958 | 0.002151 | −0.005468 | 0.006443 |
| 17 | 0.989993 | 0.990185 | 0.99379 | 0.985722 | −0.000191 | −0.003797 | 0.00427 |
| 18 | 0.995419 | 0.99617 | 0.997697 | 0.992813 | −0.00075 | −0.002277 | 0.002606 |
| 19 | 0.998021 | 0.99865 | 0.999229 | 0.996546 | −0.000628 | −0.001207 | 0.001476 |
| 20 | 0.999192 | 0.999571 | 0.999767 | 0.998412 | −0.000378 | −0.000574 | 0.000781 |
| 21 | 0.999688 | 0.999877 | 0.999937 | 0.9993 | −0.000189 | −0.000248 | 0.000388 |
| 22 | 0.999886 | 0.999968 | 0.999985 | 0.999704 | −0.000082 | −0.000098 | 0.000182 |
| 23 | 0.99996 | 0.999993 | 0.999997 | 0.99988 | −0.000032 | −0.000036 | 0.00008 |
| 24 | 0.999987 | 0.999998 | 0.999999 | 0.999953 | −0.000011 | −0.000012 | 0.000034 |
| 25 | 0.999996 | 1.0 | 1.0 | 0.999982 | −0.000003 | −0.000003 | 0.000014 |

| | $\mathcal{B}(100\ ,\ 0.1)$ | AN | ANCC | $\mathcal{P}(10)$ |
|---|---|---|---|---|
| $P([8\ ,\ 12]) = F(12) - F(7)$ | 0.59577 | 0.588852 | 0.595343 | 0.571336 |
| $P([7\ ,\ 13]) = F(13) - F(6)$ | 0.758968 | 0.750134 | 0.756655 | 0.734323 |
| $P([6\ ,\ 14]) = F(14) - F(5)$ | 0.86985 | 0.860998 | 0.866386 | 0.849456 |
| $P([5\ ,\ 15]) = F(15) - F(4)$ | 0.936398 | 0.92946 | 0.933247 | 0.922007 |
| $P([4\ ,\ 16]) = F(16) - F(3)$ | 0.971565 | 0.967435 | 0.96974 | 0.962622 |
| $P([3\ ,\ 17]) = F(17) - F(2)$ | 0.988048 | 0.986354 | 0.987581 | 0.982953 |
| $P([2\ ,\ 18]) = F(18) - F(1)$ | 0.995098 | 0.99482 | 0.995393 | 0.992314 |

$n = 100$

$p = 0.01$ , $np = 1$ , $\gamma_{100}(0.01) = 0.78811$ , $\rho_{100}(1) = 0.02$ .

| $k$ | $F$ | $\Psi$ | $\Psi^*$ | $\Pi$ | $F - \Psi$ | $F - \Psi^*$ | $F - \Pi$ |
|---|---|---|---|---|---|---|---|
| 0 | 0.366032 | 0.157439 | 0.307651 | 0.367879 | 0.208593 | 0.058381 | −0.001847 |
| **1** | 0.735762 | 0.5 | 0.692349 | 0.735759 | 0.235762 | 0.043413 | 0.000003 |
| 2 | 0.920627 | 0.842561 | 0.934166 | 0.919699 | 0.078066 | −0.013539 | 0.000928 |
| 3 | 0.981626 | 0.977788 | 0.994008 | 0.981012 | 0.003838 | −0.012381 | 0.000614 |
| 4 | 0.996568 | 0.998716 | 0.999782 | 0.99634 | −0.002147 | −0.003214 | 0.000228 |
| 5 | 0.999465 | 0.999971 | 0.999997 | 0.999406 | −0.000505 | −0.000531 | 0.00006 |
| 6 | 0.999929 | 1.0 | 1.0 | 0.999917 | −0.00007 | −0.000071 | 0.000012 |
| 7 | 0.999992 | 1.0 | 1.0 | 0.99999 | −0.000008 | −0.000008 | 0.000002 |

| | $\mathcal{B}(100\ ,\ 0.01)$ | AN | ANCC | $\mathcal{P}(1)$ |
|---|---|---|---|---|
| $P([0\ ,\ 1]) = F(1) - F(-1)$ | 0.735762 | 0.477788 | 0.626515 | 0.735759 |
| $P([0\ ,\ 2]) = F(2) - F(-1)$ | 0.920627 | 0.820349 | 0.868332 | 0.919699 |

$n = 100$

$np = 0.1$ ,

$p = 0.001$ , $\gamma_{100}(0.001) = 2.52603$ , $\rho_{100}(0.1) = 0.0002$ .

| $k$ | $F$ | $\Psi$ | $\Psi^*$ | $\Pi$ | $F - \Psi$ | $F - \Psi^*$ | $F - \Pi$ |
|---|---|---|---|---|---|---|---|
| **0** | 0.904792 | 0.375855 | 0.897162 | 0.904837 | 0.528937 | 0.00763 | −0.000045 |
| **1** | 0.995362 | 0.997797 | 0.999995 | 0.995321 | −0.002434 | −0.004633 | 0.000041 |

| | $\mathcal{B}(100\ ,\ 0.001)$ | AN | ANCC | $\mathcal{P}(0.1)$ |
|---|---|---|---|---|
| $P([0\ ,\ 1]) = F(1) - F(-1)$ | 0.995362 | 0.997546 | 0.971168 | 0.995321 |

$n = 1000$

$p = 0.01$ , $np = 10$ , $\gamma_{1000}(0.01) = 0.249222$ , $\rho_{1000}(10) = 0.04$ .

| $k$ | $F$ | $\Psi$ | $\Psi^*$ | $\Pi$ | $F-\Psi$ | $F-\Psi^*$ | $F-\Pi$ |
|---|---|---|---|---|---|---|---|
| 0 | 0.000043 | 0.000741 | 0.001267 | 0.000045 | −0.000697 | −0.001223 | −0.000002 |
| 1 | 0.000479 | 0.002116 | 0.003452 | 0.000499 | −0.001636 | −0.002972 | −0.00002 |
| 2 | 0.002679 | 0.005502 | 0.008571 | 0.002769 | −0.002822 | −0.005891 | −0.000089 |
| 3 | 0.010073 | 0.013049 | 0.019422 | 0.010336 | −0.002976 | −0.009349 | −0.000263 |
| 4 | 0.028686 | 0.028265 | 0.040231 | 0.029253 | 0.000421 | −0.011544 | −0.000566 |
| 5 | 0.06614 | 0.056018 | 0.076331 | 0.067086 | 0.010121 | −0.010191 | −0.000946 |
| 6 | 0.128877 | 0.101814 | 0.132989 | 0.130141 | 0.027063 | −0.004112 | −0.001264 |
| 7 | 0.218863 | 0.170178 | 0.213437 | 0.220221 | 0.048685 | 0.005426 | −0.001357 |
| 8 | 0.331687 | 0.262505 | 0.316777 | 0.33282 | 0.069182 | 0.014911 | −0.001132 |
| 9 | 0.457301 | 0.37531 | 0.43687 | 0.45793 | 0.08199 | 0.020431 | −0.000629 |
| **10** | 0.583041 | 0.5 | 0.56313 | 0.58304 | 0.083041 | 0.019911 | 0.000001 |
| 11 | 0.69735 | 0.62469 | 0.683223 | 0.696776 | 0.07266 | 0.014127 | 0.000574 |
| 12 | 0.792512 | 0.737495 | 0.786563 | 0.791556 | 0.055017 | 0.005949 | 0.000955 |
| 13 | 0.865565 | 0.829822 | 0.867011 | 0.864464 | 0.035743 | −0.001446 | 0.0011 |
| 14 | 0.917588 | 0.898186 | 0.923669 | 0.916542 | 0.019402 | −0.006081 | 0.001046 |
| 15 | 0.952129 | 0.943982 | 0.959769 | 0.95126 | 0.008148 | −0.007639 | 0.00087 |
| 16 | 0.973609 | 0.971735 | 0.980578 | 0.972958 | 0.001874 | −0.006969 | 0.000651 |
| 17 | 0.986167 | 0.986951 | 0.991429 | 0.985722 | −0.000783 | −0.005261 | 0.000445 |
| 18 | 0.993095 | 0.994498 | 0.996548 | 0.992813 | −0.001402 | −0.003453 | 0.000282 |
| 19 | 0.996712 | 0.997884 | 0.998733 | 0.996546 | −0.001172 | −0.002021 | 0.000166 |
| 20 | 0.998504 | 0.999259 | 0.999577 | 0.998412 | −0.000755 | −0.001073 | 0.000092 |
| 21 | 0.999348 | 0.999764 | 0.999871 | 0.9993 | −0.000415 | −0.000523 | 0.000048 |
| 22 | 0.999728 | 0.999932 | 0.999964 | 0.999704 | −0.000203 | −0.000236 | 0.000024 |
| 23 | 0.999891 | 0.999982 | 0.999991 | 0.99988 | −0.000091 | −0.0001 | 0.000011 |
| 24 | 0.999958 | 0.999996 | 0.999998 | 0.999953 | −0.000037 | −0.00004 | 0.000005 |
| 25 | 0.999984 | 0.999999 | 1.0 | 0.999982 | −0.000014 | −0.000015 | 0.000002 |

| | $\mathcal{B}(1000\ ,\ 0.01)$ | AN | ANCC | $\mathcal{P}(10)$ |
|---|---|---|---|---|
| $P([8\ ,\ 12]) = F(12) - F(7)$ | 0.573648 | 0.567317 | 0.573126 | 0.571336 |
| $P([7\ ,\ 13]) = F(13) - F(6)$ | 0.736688 | 0.728008 | 0.734022 | 0.734323 |
| $P([6\ ,\ 14]) = F(14) - F(5)$ | 0.851448 | 0.842168 | 0.847339 | 0.849456 |
| $P([5\ ,\ 15]) = F(15) - F(4)$ | 0.923443 | 0.915716 | 0.919539 | 0.922007 |
| $P([4\ ,\ 16]) = F(16) - F(3)$ | 0.963536 | 0.958686 | 0.961156 | 0.962622 |
| $P([3\ ,\ 17]) = F(17) - F(2)$ | 0.983488 | 0.981449 | 0.982858 | 0.982953 |
| $P([2\ ,\ 18]) = F(18) - F(1)$ | 0.992616 | 0.992382 | 0.993097 | 0.992314 |
| $P([1\ ,\ 19]) = F(19) - F(0)$ | 0.996668 | 0.997143 | 0.997466 | 0.9965 |

$n = 1000$

$p = 0.001$ , $np = 1$ , $\gamma_{1000}(0.001) = 0.798801$ , $\rho_{1000}(1) = 0.002$ .

| $k$ | $F$ | $\Psi$ | $\Psi^*$ | $\Pi$ | $F - \Psi$ | $F - \Psi^*$ | $F - \Pi$ |
|---|---|---|---|---|---|---|---|
| 0 | 0.367695 | 0.158534 | 0.308449 | 0.367879 | 0.209161 | 0.059246 | −0.000184 |
| **1** | 0.735759 | 0.5 | 0.691551 | 0.735759 | 0.235759 | 0.044208 | 0 |
| 2 | 0.919791 | 0.841466 | 0.93329 | 0.919699 | 0.078325 | −0.013499 | 0.000092 |
| 3 | 0.981073 | 0.977304 | 0.993812 | 0.981012 | 0.003769 | −0.012739 | 0.000061 |
| 4 | 0.996363 | 0.998657 | 0.999769 | 0.99634 | −0.002293 | −0.003405 | 0.000023 |
| 5 | 0.999412 | 0.999969 | 0.999997 | 0.999406 | −0.000556 | −0.000584 | 0.000006 |
| 6 | 0.999918 | 1.0 | 1.0 | 0.999917 | −0.000081 | −0.000081 | 0.000001 |
| 7 | 0.99999 | 1.0 | 1.0 | 0.99999 | −0.00001 | −0.00001 | 0 |

| | $\mathcal{B}(1000\ ,\ 0.001)$ | AN | ANCC | $\mathcal{P}(1)$ |
|---|---|---|---|---|
| $P([0\ ,\ 1]) = F(1) - F(-1)$ | 0.735759 | 0.477304 | 0.62484 | 0.735759 |
| $P([0\ ,\ 2]) = F(2) - F(-1)$ | 0.919791 | 0.81877 | 0.86658 | 0.919699 |

$n = 10000$

$p = 0.001$ , $np = 10$ , $\gamma_{10000}(0.001) = 0.252603$ , $\rho_{10000}(10) = 0.004$ .

| $k$ | $F$ | $\Psi$ | $\Psi^*$ | $\Pi$ | $F - \Psi$ | $F - \Psi^*$ | $F - \Pi$ |
|---|---|---|---|---|---|---|---|
| 0 | 0.000045 | 0.000778 | 0.001325 | 0.000045 | −0.000733 | −0.001279 | 0 |
| 1 | 0.000497 | 0.002203 | 0.00358 | 0.000499 | −0.001706 | −0.003082 | −0.000002 |
| 2 | 0.00276 | 0.005685 | 0.008825 | 0.002769 | −0.002925 | −0.006064 | −0.000009 |
| 3 | 0.01031 | 0.01339 | 0.019867 | 0.010336 | −0.00308 | −0.009557 | −0.000026 |
| 4 | 0.029196 | 0.028827 | 0.040919 | 0.029253 | 0.000369 | −0.011722 | −0.000056 |
| 5 | 0.066991 | 0.056833 | 0.077261 | 0.067086 | 0.010159 | −0.010269 | −0.000094 |
| 6 | 0.130015 | 0.102838 | 0.134071 | 0.130141 | 0.027177 | −0.004055 | −0.000126 |
| 7 | 0.220085 | 0.17127 | 0.214482 | 0.220221 | 0.048815 | 0.005603 | −0.000135 |
| 8 | 0.332707 | 0.263441 | 0.317544 | 0.33282 | 0.069266 | 0.015163 | −0.000112 |
| 9 | 0.457867 | 0.375855 | 0.437152 | 0.45793 | 0.082012 | 0.020715 | −0.000062 |
| **10** | 0.58304 | 0.5 | 0.562848 | 0.58304 | 0.08304 | 0.020192 | 0 |
| 11 | 0.696833 | 0.624145 | 0.682456 | 0.696776 | 0.072688 | 0.014377 | 0.000057 |
| 12 | 0.791651 | 0.736559 | 0.785518 | 0.791556 | 0.055093 | 0.006134 | 0.000095 |
| 13 | 0.864574 | 0.82873 | 0.865929 | 0.864464 | 0.035844 | −0.001355 | 0.000109 |
| 14 | 0.916646 | 0.897162 | 0.922739 | 0.916542 | 0.019484 | −0.006092 | 0.000104 |
| 15 | 0.951346 | 0.943167 | 0.959081 | 0.95126 | 0.008179 | −0.007734 | 0.000087 |
| 16 | 0.973023 | 0.971173 | 0.980133 | 0.972958 | 0.001851 | −0.007109 | 0.000065 |
| 17 | 0.985767 | 0.98661 | 0.991175 | 0.985722 | −0.000842 | −0.005408 | 0.000045 |
| 18 | 0.992842 | 0.994315 | 0.99642 | 0.992813 | −0.001472 | −0.003577 | 0.000028 |
| 19 | 0.996562 | 0.997797 | 0.998675 | 0.996546 | −0.001234 | −0.002112 | 0.000017 |
| 20 | 0.998421 | 0.999222 | 0.999553 | 0.998412 | −0.0008 | −0.001132 | 0.000009 |

| | $\mathcal{B}(10000\ ,\ 0.001)$ | AN | ANCC | $\mathcal{P}(10)$ |
|---|---|---|---|---|
| $P([8\ ,\ 12]) = F(12) - F(7)$ | 0.571566 | 0.565289 | 0.571036 | 0.571336 |
| $P([7\ ,\ 13]) = F(13) - F(6)$ | 0.734559 | 0.725892 | 0.731858 | 0.734323 |
| $P([6\ ,\ 14]) = F(14) - F(5)$ | 0.849654 | 0.840329 | 0.845477 | 0.849456 |
| $P([5\ ,\ 15]) = F(15) - F(4)$ | 0.92215 | 0.91434 | 0.918163 | 0.922007 |
| $P([4\ ,\ 16]) = F(16) - F(3)$ | 0.962714 | 0.957782 | 0.960267 | 0.962622 |
| $P([3\ ,\ 17]) = F(17) - F(2)$ | 0.983007 | 0.980924 | 0.982351 | 0.982953 |
| $P([2\ ,\ 18]) = F(18) - F(1)$ | 0.992344 | 0.992111 | 0.992839 | 0.992314 |
| $P([1\ ,\ 19]) = F(19) - F(0)$ | 0.996517 | 0.997018 | 0.99735 | 0.9965 |

$$n = 10000$$

$$p = 0.0001\ ,\quad np = 1\ ,\quad \gamma_{10000}(0.0001) = 0.79988\ ,\ \rho_{10000}(1) = 0.0002\ .$$

| $k$ | $F$ | $\Psi$ | $\Psi^*$ | $\Pi$ | $F - \Psi$ | $F - \Psi^*$ | $F - \Pi$ |
|---|---|---|---|---|---|---|---|
| 0 | 0.367861 | 0.158643 | 0.308529 | 0.367879 | 0.209218 | 0.059332 | −0.000018 |
| 1 | 0.735759 | 0.5 | 0.691471 | 0.735759 | 0.235759 | 0.044288 | 0 |
| 2 | 0.919708 | 0.841357 | 0.933203 | 0.919699 | 0.078351 | −0.013494 | 0.000009 |
| 3 | 0.981018 | 0.977255 | 0.993793 | 0.981012 | 0.003763 | −0.012774 | 0.000006 |
| 4 | 0.996342 | 0.998651 | 0.999768 | 0.99634 | −0.002308 | −0.003425 | 0.000002 |
| 5 | 0.999406 | 0.999968 | 0.999997 | 0.999406 | −0.000561 | −0.00059 | 0.000001 |
| 6 | 0.999917 | 1.0 | 1.0 | 0.999917 | −0.000082 | −0.000083 | 0 |

| | $\mathcal{B}(10000\ ,\ 0.0001)$ | AN | ANCC | $\mathcal{P}(1)$ |
|---|---|---|---|---|
| $P([0\ ,\ 2]) = F(2) - F(-1)$ | 0.919708 | 0.818612 | 0.866405 | 0.919699 |
| $P([0\ ,\ 3]) = F(3) - F(-1)$ | 0.981018 | 0.954511 | 0.926995 | 0.981012 |
| $P([0\ ,\ 4]) = F(4) - F(-1)$ | 0.996342 | 0.975906 | 0.93297 | 0.99634 |
| $P([0\ ,\ 5]) = F(5) - F(-1)$ | 0.999406 | 0.977224 | 0.933199 | 0.999406 |
| $P([0\ ,\ 6]) = F(6) - F(-1)$ | 0.999917 | 0.977255 | 0.933202 | 0.999917 |
| $P([2\ ,\ +\infty]) = 1 - F(1)$ | 0.264241 | 0.5 | 0.308529 | 0.264241 |

**Exercice 8.16.**

La désintégration d'un atome radio-actif se fait par l'émission, à des moments aléatoires, de particules énergétiques. On considère plus généralement un phénomène consistant en l'occurrence, à des moments aléatoires, d'un certain événement. On modélise ce phénomène par une famille de v.a. $(X_t)_{t\geqslant 0}$ sur un espace probabilisé $(\Omega, \mathfrak{A}, P)$, avec $X_0 = 0$. Pour tout intervalle $J = (s, t)$ de $[0, +\infty[$ ouvert, fermé ou semi-ouvert, avec $0 \leqslant s < t$, on définit la v.a. $X_J = X_t - X_s$. $X_J$ représente le nombre d'occurrences de l'événement dans l'intervalle de temps $J$. On a $X_t = X_{[0,t]} = X_{[0,t[}$ $\forall t > 0$. On fait les hypothèses suivantes (qui sont confirmées par l'expérience dans le cas de l'atome radio-actif).

(1) La loi de $X_J$ ne dépend que de la longueur de $J$ (hypothèse de *stationnarité*).

(2) Si $J_1, \cdots, J_n$ sont des intervalles deux-à-deux disjoints, les v.a. $X_{J_1}, \cdots, X_{J_n}$ sont indépendantes (hypothèse de *non-héréditarité*).

(3) Deux occurrences dans un laps de temps très court sont hautement improbables, ce qu'on formalise par la condition

$$P(X_t \geqslant 2) = o(t), \tag{8.149}$$

la notation habituelle $o(t)$ désignant une fonction telle que $\lim_{t\to 0} \frac{o(t)}{t} = 0$ (hypothèse d'*ordinarité*).

Soit $P_k(t) = P(X_t = k)$ ($k \in \mathbb{N}$, $t \geqslant 0$). On suppose $0 < P_0(1) < 1$, *i.e.* $P_0(1) = e^{-\lambda}$ avec $\lambda > 0$.
(i) Montrer que $P_k(0) = 0 \quad \forall k \geqslant 1$ et $P_0(t) = e^{-\lambda t} \quad \forall t \geqslant 0$.
(ii) Montrer que pour tous $t, h \geqslant 0$ et tout $k \geqslant 1$,

$$P_k(t+h) = P_k(t) - \lambda h P_k(t) + \lambda h P_{k-1}(t) + o(h). \tag{8.150}$$

En déduire que la fonction $t \mapsto P_k(t)$ est solution de l'équation différentielle

$$P_k'(t) = -\lambda P_k(t) + \lambda P_{k-1}(t). \tag{8.151}$$

(iii) En intégrant (8.151), montrer que $P_k(t) = e^{-\lambda t}\frac{(\lambda t)^k}{k!}$, *i.e.* la loi de $X_t$ est la loi de Poisson $\mathcal{P}(\lambda t)$ pour $t > 0$.
On dit que $(X_t)_{t\geqslant 0}$ est un *processus de Poisson* d'espérance $\lambda$ par unité de temps, ou de taux $\lambda$.
**Indication.**
(i) Soit $k \geqslant 1$. On a $P_k(0) = P(X_0 = k) = P(0 = k) = 0$. Ensuite,

$$]0,1] = \bigcup_{k=1}^{n} \left]\frac{k-1}{n}, \frac{k}{n}\right]$$

donne

$$\begin{aligned} P_0(1) &= P(X_{[0,1]} = 0) \\ &= P(X_{]0,1]} = 0) \\ &= P\left(\{X_{]0,\frac{1}{n}]} = 0\} \cap \{X_{]\frac{1}{n},\frac{2}{n}]} = 0\} \cap \cdots \cap \{X_{]\frac{n-1}{n},1]} = 0\}\right) \\ &= P\left(X_{]0,\frac{1}{n}]} = 0\right) \cdot P\left(X_{]\frac{1}{n},\frac{2}{n}]} = 0\right) \cdot \cdots \cdot P\left(X_{]\frac{n-1}{n},1]} = 0\right) \\ &= \left(P_0(\frac{1}{n})\right)^n \end{aligned}$$

donc $e^{-\lambda} = \left(P_0(\frac{1}{n})\right)^n$ ou encore $P_0(\frac{1}{n}) = e^{-\frac{\lambda}{n}}$. On obtient de même $P\left(X_{]0,\frac{k}{n}]} = 0\right) = P\left(X_{]0,\frac{1}{n}]} = 0\right) \cdot P\left(X_{]\frac{1}{n},\frac{2}{n}]} = 0\right) \cdot \cdots \cdot P\left(X_{]\frac{k-1}{n},\frac{k}{n}]} = 0\right) = \left(P_0(\frac{1}{n})\right)^k$ donc $P_0(\frac{k}{n}) = e^{-\frac{k\lambda}{n}}$.
Soit maintenant $t \geqslant 0$, $n \geqslant 1$ et $k_n$ la partie entière de $nt$. On a

$$\frac{k_n}{n} \leqslant t < \frac{k_n+1}{n}.$$

Or $P_0(t)$ est une fonction décroissante de $t$. En effet, si $s \geqslant t$,

$$\{X_{[0,s]} = 0\} \subset \{X_{[0,t]} = 0\}$$

implique $P_0(s) \leqslant P_0(t)$. Il vient donc

$$P_0(\frac{k_n}{n}) \geqslant P_0(t) \geqslant P_0(\frac{k_n+1}{n})$$

qui s'écrit

$$e^{-\lambda\frac{k_n}{n}} \geqslant P_0(t) \geqslant e^{-\lambda\frac{k_n+1}{n}}.$$

Quand $n \to +\infty$, on a $\frac{k_n}{n} \to t$ d'où

$$e^{-\lambda t} \geqslant P_0(t) \geqslant e^{-\lambda t}$$

*i.e.*

$$P_0(t) = e^{-\lambda t}.$$

(ii) On a $P_k(t+h) = P\left(X_{]0,t+h]} = k\right)$. Or

$$\left\{X_{]0,t+h]} = k\right\} = \bigcup_{j=0}^{k} \left(\left\{X_{]0,t]} = k-j\right\} \cap \left\{X_{]t,t+h]} = j\right\}\right)$$

avec une réunion disjointe. Donc

$$\begin{aligned} P_k(t+h) &= \sum_{j=0}^{k} P_{k-j}(t)\, P_j(h) \\ &= P_k(t)\, P_0(h) + P_{k-1}(t)\, P_1(h) + R(t,h) \end{aligned}$$

avec

$$R(t,h) = \sum_{j=2}^{k} P_{k-j}(t)\, P_j(h).$$

On a

$$0 \leqslant R(t,h) \leqslant \sum_{j=2}^{k} P_j(h) \leqslant P(X_h \geqslant 2) = o(h)$$

d'après (8.149) (hypothèse (3)). Donc $R(t,h)$ est aussi un $o(h)$ quand $h \to 0$. Alors $P(X_h \geqslant 2) = 1 - P_1(h) - P_0(h)$ donne

$$P_1(h) = 1 - P_0(h) + o(h) = 1 - e^{-\lambda h} + o(h) = \lambda h + o(h)$$

puisque $e^{-\lambda h} = 1 - \lambda h + o(h)$. Il en résulte

$$\begin{aligned} P_k(t+h) &= P_k(t)e^{-\lambda h} + P_{k-1}(t)P_1(h) + R(t,h) \\ &= P_k(t)(1 - \lambda h + o(h)) + P_{k-1}(t)(\lambda h + o(h)) + o(h), \end{aligned}$$

d'où (8.150). On a alors pour tout $h \neq 0$

$$\frac{1}{h}(P_k(t+h) - P_k(t)) = -\lambda P_k(t) + \lambda P_{k-1}(t) + \frac{o(h)}{h}.$$

En faisant tendre $h$ vers 0, il en résulte que la fonction $t \mapsto P_k(t)$ est dérivable et vérifie l'équation différentielle (8.151).
(iii) Posons pour tout $k \in \mathbb{N}$ $P_k(t) = C_k(t)\, e^{-\lambda t}$, *i.e.* $C_k(t) = P_k(t)\, e^{\lambda t}$. L'équation différentielle (8.151) s'écrit alors

$$C_k'(t) = \lambda\, C_{k-1}(t).$$

D'après (i), $C_0(t) = 1$ et $C_k(0) = P_k(0) = 0\ \forall k \geqslant 1$. On a alors

$$C_1'(t) = \lambda C_0(t) = \lambda \;\Rightarrow\; C_1(t) = \lambda t + \mu_1$$

avec $\mu_1 = C_1(0) = 0$ donc $C_1(t) = \lambda t$. Puis

$$C_2'(t) = \lambda C_1(t) = \lambda^2 t \Rightarrow C_2(t) = \frac{(\lambda t)^2}{2} + \mu_2$$

avec $\mu_2 = C_2(0) = 0$ donc $C_2(t) = \frac{(\lambda t)^2}{2}$. Puis par récurrence

$$C_k(t) = \frac{(\lambda t)^k}{k!}.$$

D'où $P_k(t) = e^{-\lambda t}\frac{(\lambda t)^k}{k!}$.

**Exercice 8.17.**

On suppose que les clients d'une poste arrivent suivant un processus de Poisson d'espérance 2 clients par minute. Donner (en utilisant l'ex. 8.16) un intervalle de confiance à 95% pour le nombre de clients en 2 heures.

**Indication.**

Soit $X_{120}$ le nombre de clients en 2 heures. D'après l'ex. 8.16, $X_{120}$ suit $\mathcal{P}(\lambda t)$ avec $\lambda = 2$ et $t = 120$, soit $\mathcal{P}(240)$. L'espérance et la variance de $X_{120}$ sont donc 240. Soit $\widetilde{X_{120}} = \frac{X_{120}-240}{\sqrt{240}}$ la variable centrée réduite. Avec l'approximation normale de la loi de Poisson, on a

$$P\left(-\alpha < \widetilde{X_{120}} \leqslant \alpha\right) = 0.95$$

pour $\alpha \approx 1.96$ . On en déduit l'intervalle de confiance à 95% $[210, 270]$.

**Exercice 8.18.**

(Problème des boîtes d'allumettes de Banach[1]).

Un fumeur a une boîte d'allumettes dans sa poche gauche et une autre dans sa poche droite. Chaque boîte contient initialement $N$ allumettes. Quand le fumeur veut une allumette, il sélectionne *au hasard* une poche et prend une allumette dans la boîte correspondante. Trouver la probabilité $u_r$ ($0 \leqslant r \leqslant N$) pour qu'il reste exactement $r$ allumettes dans l'autre boîte quand le fumeur découvre pour la première fois que l'une des boîtes est vide.

**Indication.**

Fixons l'une des poches par exemple la poche gauche. On peut modéliser le problème par un processus de Bernouilli infini, l'expérience élémentaire de Bernouilli consistant en le choix d'une poche et le choix de la poche gauche étant la réussite. Soit $X$ le nombre d'allumettes restant dans l'autre boîte quand le fumeur découvre pour la première fois que l'une des boîtes est vide. On a $X = r$ si et seulement si la poche que le fumeur découvre vide a été sélectionnée $N+1$ fois, et l'autre poche $N-r$ fois. Cela signifie que le temps d'attente de $N+1$ sélections de la poche découverte vide est $N+1+N-r = 2N+1-r$. L'événement $\{X = r\}$ est donc la réunion des deux événements suivants :

- $A$ : "Le temps d'attente de $N+1$ sélections de la poche gauche (réussite à l'épreuve élémentaire de Bernouilli) est $2N+1-r$".
- $B$ : "Le temps d'attente de $N+1$ sélections de la poche droite (échec à l'épreuve élémentaire de Bernouilli) est $2N+1-r$".

Les événements $A$ et $B$ sont incompatibles.

1. Le nom de ce problème est, selon [10] p.166, inspiré par une référence humoristique aux habitudes de fumeur de S. Banach dans une conférence en son honneur faite par H. Steinhaus.

Or le temps d'attente de $N+1$ succès de la sélection de la poche gauche suit la loi binomiale négative $\mathcal{B}^-(N+1,\frac{1}{2})$. Il en est de même du temps d'attente de $N+1$ succès de la sélection de la poche droite. Il en résulte que

$$\begin{aligned} u_r &= 2\binom{-N-1}{N-r}(\frac{1}{2})^{N+1}(-\frac{1}{2})^{N-r} \\ &= 2\,\frac{(-N-1)\cdots(-N-1-(N-r-1))}{(N-r)!}\;\;\frac{(-1)^{N-r}}{2^{2N-r+1}} \\ &= C_{2N-r}^{N}\,\frac{1}{2^{2N-r}}\,. \end{aligned}$$

**Exercice 8.19.**

(Approximation de la loi hypergéométrique par la loi binomiale).

Soit $(X_N)_{N\geqslant 1}$ une suite de v.a. On suppose que pour chaque $N$, $X_N$ suit une loi hypergéométrique $\mathcal{H}(N,N_1,n)$ de paramètres $N,N_1,n$ où $n$ est fixé et $N_1$ est une fonction de $N$ telle que $\lim_{N\to+\infty}\frac{N_1}{N}=p$ , $0<p<1$. On suppose aussi que $N_1\geqslant n$ et $N_2=N-N_1\geqslant n$. On sait qu'alors l'ensemble fondamental $\Omega_{(N,N_1,n)}$ est $\{0,\cdots,n\}$. Montrer que pour tout $k$, $0\leqslant k\leqslant n$ fixé,

$$P(X_N=k)\to C_n^k\,p^kq^{n-k}\quad\text{quand } N\to+\infty\,, \tag{8.152}$$

où $q=1-p$. En déduire que $X_N\Rightarrow\mathcal{B}(n,p)$ quand $N\to+\infty$ . Pour $N,N_1$ "grands", on a donc l'approximation $\mathcal{H}(N,N_1,n)\approx\mathcal{B}(n,p)$ avec $p=\frac{N_1}{N}$.

**Indication.**

On raisonne par récurrence sur la valeur de $n$. Si $n=1$, le résultat est vrai : $P(X_N=0)=\frac{N_2}{N}\to q=C_1^0p^0q^1$ et $P(X_N=1)=\frac{N_1}{N}\to p=C_1^1p^1q^0$ quand $N\to+\infty$. Supposons le résultat vrai si $n=m$ et montrons qu'il est vrai si $n=m+1$. Pour tout $k$, $1\leqslant k\leqslant m+1$ on a :

$$\begin{aligned} P(X_N=k) &= \frac{C_{N_1}^kC_{N_2}^{m+1-k}}{C_N^{m+1}}=\frac{m+1}{N-m}\;\frac{N_1-k+1}{k}\;\frac{C_{N_1}^{k-1}C_{N_2}^{m-(k-1)}}{C_N^m} \\ &\qquad(\text{car } C_{N_1}^k=\frac{N_1-k+1}{k}\,C_{N_1}^{k-1}\text{ et } C_N^{m+1}=\frac{N-m}{m+1}\,C_N^m\;) \\ &\sim \frac{m+1}{k}\,p\,\frac{C_{N_1}^{k-1}\,C_{N_2}^{m-(k-1)}}{C_N^m}\quad\text{quand } N\to+\infty \\ &\sim \frac{m+1}{k}\,p\,C_m^{k-1}\,p^{k-1}\,q^{m-(k-1)} \\ &\qquad(\text{hypothèse de récurrence appliquée à } m \text{ et } k-1,\ 0\leqslant k-1\leqslant m) \\ &\sim C_{m+1}^k\,p^kq^{m+1-k}\quad(\text{car } C_{m+1}^k=\frac{m+1}{k}\,C_m^{k-1}). \end{aligned}$$

Pour $k=0$, on obtient de même

$$P(X_N=0)=\frac{m+1}{N-m}\,\frac{C_{N_2}^{m+1}}{C_N^m}=\frac{N_2-m}{N-m}\,\frac{C_{N_2}^m}{C_N^m}\sim(1-p)\,C_m^0\,p^0q^m=C_{m+1}^0\,p^0q^{m+1}.$$

D'où (8.152). Maintenant, pour tout $t\in\mathbb{R}$,

$$P(X_N\leqslant t)=\sum_{0\leqslant k\leqslant t}P(X_N=k)\to\sum_{0\leqslant k\leqslant t}C_n^kp^kq^{n-k}=F(t)\quad\text{quand } N\to+\infty$$

où $F$ désigne la fonction de répartition de la loi binomiale $\mathcal{B}(n,p)$. Donc $X_N\Rightarrow\mathcal{B}(n,p)$ quand $N\to+\infty$.

Exemples de valeurs de la fonction de répartition $H(N, N_1, n)$ de $\mathcal{H}(N, N_1, n)$ et de la fonction de répartition $F$ de $\mathcal{B}(n, p)$, $p = \frac{N_1}{N}$ :

$$F(k) = \sum_{j=0}^{k} C_n^j p^j (1-p)^{n-j},$$

$$H(N, N_1, n)(k) = \sum_{j=0}^{k} \frac{C_{N_1}^j C_{N_2}^{n-j}}{C_N^n} \; (N_1 \geqslant n, N_2 = N - N_1 \geqslant n).$$

| $k$ | $\mathcal{B}(20, 0.25)$ | $H(100, 25, 20)$ | $H(300, 75, 20)$ | $H(1200, 300, 20)$ |
|---|---|---|---|---|
| 0 | 0.003171 | 0.001498 | 0.002539 | 0.003006 |
| 1 | 0.024313 | 0.014878 | 0.021025 | 0.023479 |
| 2 | 0.09126 | 0.068396 | 0.083807 | 0.089414 |
| 3 | 0.225156 | 0.195731 | 0.21601 | 0.222925 |
| 4 | 0.414842 | 0.397524 | 0.409571 | 0.413564 |
| 5 | 0.617173 | 0.623533 | 0.618986 | 0.617602 |
| 6 | 0.785782 | 0.808787 | 0.792671 | 0.787441 |
| 7 | 0.898188 | 0.922329 | 0.90573 | 0.900029 |
| 8 | 0.959075 | 0.975045 | 0.964383 | 0.960396 |
| 9 | 0.986136 | 0.993715 | 0.988867 | 0.986834 |
| 10 | 0.996058 | 0.998771 | 0.997135 | 0.996342 |
| 11 | 0.999065 | 0.999815 | 0.999397 | 0.999156 |
| 12 | 0.999816 | 0.999979 | 0.999897 | 0.999839 |
| 13 | 0.99997 | 0.999998 | 0.999986 | 0.999975 |
| 14 | 0.999996 | 1.0 | 0.999998 | 0.999997 |
| 15 | 1.0 | 1.0 | 1.0 | 1.0 |

| $k$ | $\mathcal{B}(9, 0.1)$ | $H(100, 10, 9)$ | $H(500, 50, 9)$ | $H(1000, 100, 9)$ |
|---|---|---|---|---|
| 0 | 0.38742 | 0.371276 | 0.384296 | 0.385865 |
| 1 | 0.774841 | 0.778774 | 0.775548 | 0.77519 |
| 2 | 0.947028 | 0.95552 | 0.948653 | 0.947836 |
| 3 | 0.991669 | 0.994796 | 0.992319 | 0.991995 |
| 4 | 0.999109 | 0.999648 | 0.999237 | 0.999174 |
| 5 | 0.999936 | 0.999987 | 0.99995 | 0.999943 |
| 6 | 0.999997 | 1.0 | 0.999998 | 0.999997 |
| 7 | 1.0 | 1.0 | 1.0 | 1.0 |

**Exercice 8.20.**

Au *"Loto"*, 6 numéros "gagnants" sont tirés sans remise parmi 49 . On effectue un pronostic sur les 6 numéros gagnants avant tirage. Quelle est la loi de la v.a. "$X$ = nombre de numéros gagnants dans le pronostic" ?

**Indication.**

$X$ suit $\mathcal{H}(49, 6, 6)$. On a $E(X) = 6 \cdot \frac{6}{49} \approx 0.73$. Les valeurs de $P(X = k)$ et celles obtenues avec l'approximation binomiale $\mathcal{B}(6, \frac{6}{49})$ sont données ci-dessous.

| $k$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $\mathcal{H}(49, 6, 6)$ | 0.435965 | 0.413019 | 0.132378 | 0.0176504 | 0.00096862 | 0.0000184499 | 0.0000000715 |
| $\mathcal{B}(6, \frac{6}{49})$ | 0.456703 | 0.382356 | 0.13338 | 0.0248149 | 0.00259691 | 0.000144944 | 0.00000337078 |

# TABLES

Table 1 : Loi normale centrée réduite

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} \, dt$$

| $x$ | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0.004 | 0.008 | 0.012 | 0.016 | 0.0199 | 0.0239 | 0.0279 | 0.0319 | 0.0359 |
| 0.1 | 0.0398 | 0.0438 | 0.0478 | 0.0517 | 0.0557 | 0.0596 | 0.0636 | 0.0675 | 0.0714 | 0.0753 |
| 0.2 | 0.0793 | 0.0832 | 0.0871 | 0.091 | 0.0948 | 0.0987 | 0.1026 | 0.1064 | 0.1103 | 0.1141 |
| 0.3 | 0.1179 | 0.1217 | 0.1255 | 0.1293 | 0.1331 | 0.1368 | 0.1406 | 0.1443 | 0.148 | 0.1517 |
| 0.4 | 0.1554 | 0.1591 | 0.1628 | 0.1664 | 0.17 | 0.1736 | 0.1772 | 0.1808 | 0.1844 | 0.1879 |
| 0.5 | 0.1915 | 0.195 | 0.1985 | 0.2019 | 0.2054 | 0.2088 | 0.2123 | 0.2157 | 0.219 | 0.2224 |
| 0.6 | 0.2257 | 0.2291 | 0.2324 | 0.2357 | 0.2389 | 0.2422 | 0.2454 | 0.2486 | 0.2517 | 0.2549 |
| 0.7 | 0.258 | 0.2611 | 0.2642 | 0.2673 | 0.2704 | 0.2734 | 0.2764 | 0.2794 | 0.2823 | 0.2852 |
| 0.8 | 0.2881 | 0.291 | 0.2939 | 0.2967 | 0.2995 | 0.3023 | 0.3051 | 0.3078 | 0.3106 | 0.3133 |
| 0.9 | 0.3159 | 0.3186 | 0.3212 | 0.3238 | 0.3264 | 0.3289 | 0.3315 | 0.334 | 0.3365 | 0.3389 |
| 1.0 | 0.3413 | 0.3438 | 0.3461 | 0.3485 | 0.3508 | 0.3531 | 0.3554 | 0.3577 | 0.3599 | 0.3621 |
| 1.1 | 0.3643 | 0.3665 | 0.3686 | 0.3708 | 0.3729 | 0.3749 | 0.377 | 0.379 | 0.381 | 0.383 |
| 1.2 | 0.3849 | 0.3869 | 0.3888 | 0.3907 | 0.3925 | 0.3944 | 0.3962 | 0.398 | 0.3997 | 0.4015 |
| 1.3 | 0.4032 | 0.4049 | 0.4066 | 0.4082 | 0.4099 | 0.4115 | 0.4131 | 0.4147 | 0.4162 | 0.4177 |
| 1.4 | 0.4192 | 0.4207 | 0.4222 | 0.4236 | 0.4251 | 0.4265 | 0.4279 | 0.4292 | 0.4306 | 0.4319 |
| 1.5 | 0.4332 | 0.4345 | 0.4357 | 0.437 | 0.4382 | 0.4394 | 0.4406 | 0.4418 | 0.4429 | 0.4441 |
| 1.6 | 0.4452 | 0.4463 | 0.4474 | 0.4484 | 0.4495 | 0.4505 | 0.4515 | 0.4525 | 0.4535 | 0.4545 |
| 1.7 | 0.4554 | 0.4564 | 0.4573 | 0.4582 | 0.4591 | 0.4599 | 0.4608 | 0.4616 | 0.4625 | 0.4633 |
| 1.8 | 0.4641 | 0.4649 | 0.4656 | 0.4664 | 0.4671 | 0.4678 | 0.4686 | 0.4693 | 0.4699 | 0.4706 |
| 1.9 | 0.4713 | 0.4719 | 0.4726 | 0.4732 | 0.4738 | 0.4744 | 0.475 | 0.4756 | 0.4761 | 0.4767 |
| 2.0 | 0.4772 | 0.4778 | 0.4783 | 0.4788 | 0.4793 | 0.4798 | 0.4803 | 0.4808 | 0.4812 | 0.4817 |
| 2.1 | 0.4821 | 0.4826 | 0.483 | 0.4834 | 0.4838 | 0.4842 | 0.4846 | 0.485 | 0.4854 | 0.4857 |
| 2.2 | 0.4861 | 0.4864 | 0.4868 | 0.4871 | 0.4875 | 0.4878 | 0.4881 | 0.4884 | 0.4887 | 0.489 |
| 2.3 | 0.4893 | 0.4896 | 0.4898 | 0.4901 | 0.4904 | 0.4906 | 0.4909 | 0.4911 | 0.4913 | 0.4916 |
| 2.4 | 0.4918 | 0.492 | 0.4922 | 0.4925 | 0.4927 | 0.4929 | 0.4931 | 0.4932 | 0.4934 | 0.4936 |
| 2.5 | 0.4938 | 0.494 | 0.4941 | 0.4943 | 0.4945 | 0.4946 | 0.4948 | 0.4949 | 0.4951 | 0.4952 |
| 2.6 | 0.4953 | 0.4955 | 0.4956 | 0.4957 | 0.4959 | 0.496 | 0.4961 | 0.4962 | 0.4963 | 0.4964 |
| 2.7 | 0.4965 | 0.4966 | 0.4967 | 0.4968 | 0.4969 | 0.497 | 0.4971 | 0.4972 | 0.4973 | 0.4974 |
| 2.8 | 0.4974 | 0.4975 | 0.4976 | 0.4977 | 0.4977 | 0.4978 | 0.4979 | 0.4979 | 0.498 | 0.4981 |
| 2.9 | 0.4981 | 0.4982 | 0.4982 | 0.4983 | 0.4984 | 0.4984 | 0.4985 | 0.4985 | 0.4986 | 0.4986 |
| 3.0 | 0.4987 | 0.4987 | 0.4987 | 0.4988 | 0.4988 | 0.4989 | 0.4989 | 0.4989 | 0.499 | 0.499 |
| 3.1 | 0.499 | 0.4991 | 0.4991 | 0.4991 | 0.4992 | 0.4992 | 0.4992 | 0.4992 | 0.4993 | 0.4993 |
| 3.2 | 0.4993 | 0.4993 | 0.4994 | 0.4994 | 0.4994 | 0.4994 | 0.4994 | 0.4995 | 0.4995 | 0.4995 |
| 3.3 | 0.4995 | 0.4995 | 0.4995 | 0.4996 | 0.4996 | 0.4996 | 0.4996 | 0.4996 | 0.4996 | 0.4997 |
| 3.4 | 0.4997 | 0.4997 | 0.4997 | 0.4997 | 0.4997 | 0.4997 | 0.4997 | 0.4997 | 0.4997 | 0.4998 |
| 3.5 | 0.4998 | 0.4998 | 0.4998 | 0.4998 | 0.4998 | 0.4998 | 0.4998 | 0.4998 | 0.4998 | 0.4998 |
| 3.6 | 0.4998 | 0.4998 | 0.4999 | 0.4999 | 0.4999 | 0.4999 | 0.4999 | 0.4999 | 0.4999 | 0.4999 |

Table 2 : Fonction de répartition $F$ de la loi binomiale $\mathcal{B}(n, p)$ :

$$F(s) = \sum_{k=0}^{s} C_n^k p^k (1-p)^{n-k}.$$

| $n$ | $s \setminus p$ | 0.01 | 0.05 | 0.1 | 0.2 | 0.25 | 0.3 | 0.4 | 0.5 |
|---|---|---|---|---|---|---|---|---|---|
| 5 | 0 | 0.951 | 0.7738 | 0.5905 | 0.3277 | 0.2373 | 0.1681 | 0.0778 | 0.0312 |
| | 1 | 0.999 | 0.9774 | 0.9185 | 0.7373 | 0.6328 | 0.5282 | 0.337 | 0.1875 |
| | 2 | 1.0 | 0.9988 | 0.9914 | 0.9421 | 0.8965 | 0.8369 | 0.6826 | 0.5 |
| | 3 | 1.0 | 1.0 | 0.9995 | 0.9933 | 0.9844 | 0.9692 | 0.913 | 0.8125 |
| | 4 | 1.0 | 1.0 | 1.0 | 0.9997 | 0.999 | 0.9976 | 0.9898 | 0.9687 |
| | 5 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |

| $n$ | $s \setminus p$ | 0.01 | 0.05 | 0.1 | 0.2 | 0.25 | 0.3 | 0.4 | 0.5 |
|---|---|---|---|---|---|---|---|---|---|
| 10 | 0 | 0.9044 | 0.5987 | 0.3487 | 0.1074 | 0.0563 | 0.0282 | 0.006 | 0.001 |
| | 1 | 0.9957 | 0.9139 | 0.7361 | 0.3758 | 0.244 | 0.1493 | 0.0464 | 0.0107 |
| | 2 | 0.9999 | 0.9885 | 0.9298 | 0.6778 | 0.5256 | 0.3828 | 0.1673 | 0.0547 |
| | 3 | 1.0 | 0.999 | 0.9872 | 0.8791 | 0.7759 | 0.6496 | 0.3823 | 0.1719 |
| | 4 | 1.0 | 0.9999 | 0.9984 | 0.9672 | 0.9219 | 0.8497 | 0.6331 | 0.377 |
| | 5 | 1.0 | 1.0 | 0.9999 | 0.9936 | 0.9803 | 0.9527 | 0.8338 | 0.623 |
| | 6 | 1.0 | 1.0 | 1.0 | 0.9991 | 0.9965 | 0.9894 | 0.9452 | 0.8281 |
| | 7 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.9996 | 0.9984 | 0.9877 | 0.9453 |
| | 8 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.9983 | 0.9893 |
| | 9 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.999 |
| | 10 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |

| $n$ | $s \setminus p$ | 0.01 | 0.05 | 0.1 | 0.2 | 0.25 | 0.3 | 0.4 | 0.5 |
|---|---|---|---|---|---|---|---|---|---|
| 15 | 0 | 0.8601 | 0.4633 | 0.2059 | 0.0352 | 0.0134 | 0.0047 | 0.0005 | 0 |
| | 1 | 0.9904 | 0.829 | 0.549 | 0.1671 | 0.0802 | 0.0353 | 0.0052 | 0.0005 |
| | 2 | 0.9996 | 0.9638 | 0.8159 | 0.398 | 0.2361 | 0.1268 | 0.0271 | 0.0037 |
| | 3 | 1.0 | 0.9945 | 0.9444 | 0.6482 | 0.4613 | 0.2969 | 0.0905 | 0.0176 |
| | 4 | 1.0 | 0.9994 | 0.9873 | 0.8358 | 0.6865 | 0.5155 | 0.2173 | 0.0592 |
| | 5 | 1.0 | 0.9999 | 0.9978 | 0.9389 | 0.8516 | 0.7216 | 0.4032 | 0.1509 |
| | 6 | 1.0 | 1.0 | 0.9997 | 0.9819 | 0.9434 | 0.8689 | 0.6098 | 0.3036 |
| | 7 | 1.0 | 1.0 | 1.0 | 0.9958 | 0.9827 | 0.95 | 0.7869 | 0.5 |
| | 8 | 1.0 | 1.0 | 1.0 | 0.9992 | 0.9958 | 0.9848 | 0.905 | 0.6964 |
| | 9 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.9992 | 0.9963 | 0.9662 | 0.8491 |
| | 10 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.9993 | 0.9907 | 0.9408 |
| | 11 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.9981 | 0.9824 |
| | 12 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9997 | 0.9963 |
| | 13 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9995 |
| | 14 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |

| $n$ | $s \setminus p$ | 0.01 | 0.05 | 0.1 | 0.2 | 0.25 | 0.3 | 0.4 | 0.5 |
|---|---|---|---|---|---|---|---|---|---|
| 20 | 0 | 0.8179 | 0.3585 | 0.1216 | 0.0115 | 0.0032 | 0.0008 | 0 | 0 |
| | 1 | 0.9831 | 0.7358 | 0.3917 | 0.0692 | 0.0243 | 0.0076 | 0.0005 | 0 |
| | 2 | 0.999 | 0.9245 | 0.6769 | 0.2061 | 0.0913 | 0.0355 | 0.0036 | 0.0002 |
| | 3 | 1.0 | 0.9841 | 0.867 | 0.4114 | 0.2252 | 0.1071 | 0.016 | 0.0013 |
| | 4 | 1.0 | 0.9974 | 0.9568 | 0.6296 | 0.4148 | 0.2375 | 0.051 | 0.0059 |
| | 5 | 1.0 | 0.9997 | 0.9887 | 0.8042 | 0.6172 | 0.4164 | 0.1256 | 0.0207 |
| | 6 | 1.0 | 1.0 | 0.9976 | 0.9133 | 0.7858 | 0.608 | 0.25 | 0.0577 |
| | 7 | 1.0 | 1.0 | 0.9996 | 0.9679 | 0.8982 | 0.7723 | 0.4159 | 0.1316 |
| | 8 | 1.0 | 1.0 | 0.9999 | 0.99 | 0.9591 | 0.8867 | 0.5956 | 0.2517 |
| | 9 | 1.0 | 1.0 | 1.0 | 0.9974 | 0.9861 | 0.952 | 0.7553 | 0.4119 |
| | 10 | 1.0 | 1.0 | 1.0 | 0.9994 | 0.9961 | 0.9829 | 0.8725 | 0.5881 |
| | 11 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.9991 | 0.9949 | 0.9435 | 0.7483 |
| | 12 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9998 | 0.9987 | 0.979 | 0.8684 |
| | 13 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9997 | 0.9935 | 0.9423 |
| | 14 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9984 | 0.9793 |
| | 15 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9997 | 0.9941 |
| | 16 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9987 |
| | 17 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9998 |
| | 18 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |

| $n$ | $s \setminus p$ | 0.01 | 0.05 | 0.1 | 0.2 | 0.25 | 0.3 | 0.4 | 0.5 |
|---|---|---|---|---|---|---|---|---|---|
| 25 | 0 | 0.7778 | 0.2774 | 0.0718 | 0.0038 | 0.0008 | 0.0001 | 0 | 0 |
| | 1 | 0.9742 | 0.6424 | 0.2712 | 0.0274 | 0.007 | 0.0016 | 0.0001 | 0 |
| | 2 | 0.998 | 0.8729 | 0.5371 | 0.0982 | 0.0321 | 0.009 | 0.0004 | 0 |
| | 3 | 0.9999 | 0.9659 | 0.7636 | 0.234 | 0.0962 | 0.0332 | 0.0024 | 0.0001 |
| | 4 | 1.0 | 0.9928 | 0.902 | 0.4207 | 0.2137 | 0.0905 | 0.0095 | 0.0005 |
| | 5 | 1.0 | 0.9988 | 0.9666 | 0.6167 | 0.3783 | 0.1935 | 0.0294 | 0.002 |
| | 6 | 1.0 | 0.9998 | 0.9905 | 0.78 | 0.5611 | 0.3407 | 0.0736 | 0.0073 |
| | 7 | 1.0 | 1.0 | 0.9977 | 0.8909 | 0.7265 | 0.5118 | 0.1536 | 0.0216 |
| | 8 | 1.0 | 1.0 | 0.9995 | 0.9532 | 0.8506 | 0.6769 | 0.2735 | 0.0539 |
| | 9 | 1.0 | 1.0 | 0.9999 | 0.9827 | 0.9287 | 0.8106 | 0.4246 | 0.1148 |
| | 10 | 1.0 | 1.0 | 1.0 | 0.9944 | 0.9703 | 0.9022 | 0.5858 | 0.2122 |
| | 11 | 1.0 | 1.0 | 1.0 | 0.9985 | 0.9893 | 0.9558 | 0.7323 | 0.345 |
| | 12 | 1.0 | 1.0 | 1.0 | 0.9996 | 0.9966 | 0.9825 | 0.8462 | 0.5 |
| | 13 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.9991 | 0.994 | 0.9222 | 0.655 |
| | 14 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9998 | 0.9982 | 0.9656 | 0.7878 |
| | 15 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9995 | 0.9868 | 0.8852 |
| | 16 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.9957 | 0.9461 |
| | 17 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9988 | 0.9784 |
| | 18 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9997 | 0.9927 |
| | 19 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.998 |
| | 20 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9995 |
| | 21 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9999 |
| | 22 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |

| $n$ | $s \setminus p$ | 0.01 | 0.05 | 0.1 | 0.2 | 0.25 | 0.3 | 0.4 | 0.5 |
|---|---|---|---|---|---|---|---|---|---|
| 50 | 0 | 0.605 | 0.0769 | 0.0052 | 0 | 0 | 0 | 0 | 0 |
| | 1 | 0.9106 | 0.2794 | 0.0338 | 0.0002 | 0 | 0 | 0 | 0 |
| | 2 | 0.9862 | 0.5405 | 0.1117 | 0.0013 | 0.0001 | 0 | 0 | 0 |
| | 3 | 0.9984 | 0.7604 | 0.2503 | 0.0057 | 0.0005 | 0 | 0 | 0 |
| | 4 | 0.9999 | 0.8964 | 0.4312 | 0.0185 | 0.0021 | 0.0002 | 0 | 0 |
| | 5 | 1.0 | 0.9622 | 0.6161 | 0.048 | 0.007 | 0.0007 | 0 | 0 |
| | 6 | 1.0 | 0.9882 | 0.7702 | 0.1034 | 0.0194 | 0.0025 | 0 | 0 |
| | 7 | 1.0 | 0.9968 | 0.8779 | 0.1904 | 0.0453 | 0.0073 | 0.0001 | 0 |
| | 8 | 1.0 | 0.9992 | 0.9421 | 0.3073 | 0.0916 | 0.0183 | 0.0002 | 0 |
| | 9 | 1.0 | 0.9998 | 0.9755 | 0.4437 | 0.1637 | 0.0402 | 0.0008 | 0 |
| | 10 | 1.0 | 1.0 | 0.9906 | 0.5836 | 0.2622 | 0.0789 | 0.0022 | 0 |
| | 11 | 1.0 | 1.0 | 0.9968 | 0.7107 | 0.3816 | 0.139 | 0.0057 | 0 |
| | 12 | 1.0 | 1.0 | 0.999 | 0.8139 | 0.511 | 0.2229 | 0.0133 | 0.0002 |
| | 13 | 1.0 | 1.0 | 0.9997 | 0.8894 | 0.637 | 0.3279 | 0.028 | 0.0005 |
| | 14 | 1.0 | 1.0 | 0.9999 | 0.9393 | 0.7481 | 0.4468 | 0.054 | 0.0013 |
| | 15 | 1.0 | 1.0 | 1.0 | 0.9692 | 0.8369 | 0.5692 | 0.0955 | 0.0033 |
| | 16 | 1.0 | 1.0 | 1.0 | 0.9856 | 0.9017 | 0.6839 | 0.1561 | 0.0077 |
| | 17 | 1.0 | 1.0 | 1.0 | 0.9937 | 0.9449 | 0.7822 | 0.2369 | 0.0164 |
| | 18 | 1.0 | 1.0 | 1.0 | 0.9975 | 0.9713 | 0.8594 | 0.3356 | 0.0325 |
| | 19 | 1.0 | 1.0 | 1.0 | 0.9991 | 0.9861 | 0.9152 | 0.4465 | 0.0595 |
| | 20 | 1.0 | 1.0 | 1.0 | 0.9997 | 0.9937 | 0.9522 | 0.561 | 0.1013 |
| | 21 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.9974 | 0.9749 | 0.6701 | 0.1611 |
| | 22 | 1.0 | 1.0 | 1.0 | 1.0 | 0.999 | 0.9877 | 0.766 | 0.2399 |
| | 23 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9996 | 0.9944 | 0.8438 | 0.3359 |
| | 24 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.9976 | 0.9022 | 0.4439 |
| | 25 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9991 | 0.9427 | 0.5561 |
| | 26 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9997 | 0.9686 | 0.6641 |
| | 27 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.984 | 0.7601 |
| | 28 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9924 | 0.8389 |
| | 29 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9966 | 0.8987 |
| | 30 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9986 | 0.9405 |
| | 31 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9995 | 0.9675 |
| | 32 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9998 | 0.9836 |
| | 33 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.9923 |
| | 34 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9967 |
| | 35 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9987 |
| | 36 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9995 |
| | 37 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9998 |
| | 38 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |

| $n$ | $s \setminus p$ | 0.01 | 0.05 | 0.1 | 0.2 | 0.25 | 0.3 | 0.4 | 0.5 |
|---|---|---|---|---|---|---|---|---|---|
| 100 | 0 | 0.366 | 0.0059 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 1 | 0.7358 | 0.0371 | 0.0003 | 0 | 0 | 0 | 0 | 0 |
| | 2 | 0.9206 | 0.1183 | 0.0019 | 0 | 0 | 0 | 0 | 0 |
| | 3 | 0.9816 | 0.2578 | 0.0078 | 0 | 0 | 0 | 0 | 0 |
| | 4 | 0.9966 | 0.436 | 0.0237 | 0 | 0 | 0 | 0 | 0 |
| | 5 | 0.9995 | 0.616 | 0.0576 | 0 | 0 | 0 | 0 | 0 |
| | 6 | 0.9999 | 0.766 | 0.1172 | 0.0001 | 0 | 0 | 0 | 0 |
| | 7 | 1.0 | 0.872 | 0.2061 | 0.0003 | 0 | 0 | 0 | 0 |
| | 8 | 1.0 | 0.9369 | 0.3209 | 0.0009 | 0 | 0 | 0 | 0 |
| | 9 | 1.0 | 0.9718 | 0.4513 | 0.0023 | 0 | 0 | 0 | 0 |
| | 10 | 1.0 | 0.9885 | 0.5832 | 0.0057 | 0.0001 | 0 | 0 | 0 |
| | 11 | 1.0 | 0.9957 | 0.703 | 0.0126 | 0.0004 | 0 | 0 | 0 |
| | 12 | 1.0 | 0.9985 | 0.8018 | 0.0253 | 0.001 | 0 | 0 | 0 |
| | 13 | 1.0 | 0.9995 | 0.8761 | 0.0469 | 0.0025 | 0.0001 | 0 | 0 |
| | 14 | 1.0 | 0.9999 | 0.9274 | 0.0804 | 0.0054 | 0.0002 | 0 | 0 |
| | 15 | 1.0 | 1.0 | 0.9601 | 0.1285 | 0.0111 | 0.0004 | 0 | 0 |
| | 16 | 1.0 | 1.0 | 0.9794 | 0.1923 | 0.0211 | 0.001 | 0 | 0 |
| | 17 | 1.0 | 1.0 | 0.99 | 0.2712 | 0.0376 | 0.0022 | 0 | 0 |
| | 18 | 1.0 | 1.0 | 0.9954 | 0.3621 | 0.063 | 0.0045 | 0 | 0 |
| | 19 | 1.0 | 1.0 | 0.998 | 0.4602 | 0.0995 | 0.0089 | 0 | 0 |
| | 20 | 1.0 | 1.0 | 0.9992 | 0.5595 | 0.1488 | 0.0165 | 0 | 0 |
| | 21 | 1.0 | 1.0 | 0.9997 | 0.654 | 0.2114 | 0.0288 | 0 | 0 |
| | 22 | 1.0 | 1.0 | 0.9999 | 0.7389 | 0.2864 | 0.0479 | 0.0001 | 0 |
| | 23 | 1.0 | 1.0 | 1.0 | 0.8109 | 0.3711 | 0.0755 | 0.0003 | 0 |
| | 24 | 1.0 | 1.0 | 1.0 | 0.8686 | 0.4617 | 0.1136 | 0.0006 | 0 |
| | 25 | 1.0 | 1.0 | 1.0 | 0.9125 | 0.5535 | 0.1631 | 0.0012 | 0 |
| | 26 | 1.0 | 1.0 | 1.0 | 0.9442 | 0.6417 | 0.2244 | 0.0024 | 0 |
| | 27 | 1.0 | 1.0 | 1.0 | 0.9658 | 0.7224 | 0.2964 | 0.0046 | 0 |
| | 28 | 1.0 | 1.0 | 1.0 | 0.98 | 0.7925 | 0.3768 | 0.0084 | 0 |
| | 29 | 1.0 | 1.0 | 1.0 | 0.9888 | 0.8505 | 0.4623 | 0.0148 | 0 |
| | 30 | 1.0 | 1.0 | 1.0 | 0.9939 | 0.8962 | 0.5491 | 0.0248 | 0 |
| | 31 | 1.0 | 1.0 | 1.0 | 0.9969 | 0.9307 | 0.6331 | 0.0398 | 0.0001 |
| | 32 | 1.0 | 1.0 | 1.0 | 0.9984 | 0.9554 | 0.7107 | 0.0615 | 0.0002 |
| | 33 | 1.0 | 1.0 | 1.0 | 0.9993 | 0.9724 | 0.7793 | 0.0913 | 0.0004 |
| | 34 | 1.0 | 1.0 | 1.0 | 0.9997 | 0.9836 | 0.8371 | 0.1303 | 0.0009 |
| | 35 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.9906 | 0.8839 | 0.1795 | 0.0018 |
| | 36 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.9948 | 0.9201 | 0.2386 | 0.0033 |
| | 37 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9973 | 0.947 | 0.3068 | 0.006 |
| | 38 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9986 | 0.966 | 0.3822 | 0.0105 |
| | 39 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9993 | 0.979 | 0.4621 | 0.0176 |
| | 40 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9997 | 0.9875 | 0.5433 | 0.0284 |

| $n$ | $s \setminus p$ | 0.01 | 0.05 | 0.1 | 0.2 | 0.25 | 0.3 | 0.4 | 0.5 |
|---|---|---|---|---|---|---|---|---|---|
| 100 | 41 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.9928 | 0.6225 | 0.0443 |
| | 42 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.996 | 0.6967 | 0.0666 |
| | 43 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9979 | 0.7635 | 0.0967 |
| | 44 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9989 | 0.8211 | 0.1356 |
| | 45 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9995 | 0.8689 | 0.1841 |
| | 46 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9997 | 0.907 | 0.2421 |
| | 47 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.9362 | 0.3086 |
| | 48 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.9577 | 0.3822 |
| | 49 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9729 | 0.4602 |
| | 50 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9832 | 0.5398 |
| | 51 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.99 | 0.6178 |
| | 52 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9942 | 0.6914 |
| | 53 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9968 | 0.7579 |
| | 54 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9983 | 0.8159 |
| | 55 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9991 | 0.8644 |
| | 56 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9996 | 0.9033 |
| | 57 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9998 | 0.9334 |
| | 58 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.9557 |
| | 59 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9716 |
| | 60 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9824 |
| | 61 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9895 |
| | 62 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.994 |
| | 63 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9967 |
| | 64 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9982 |
| | 65 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9991 |
| | 66 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9996 |
| | 67 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9998 |
| | 68 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9999 |
| | 69 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |

Table 3 : Fonction de répartition $\Pi$ de la loi de Poisson $\mathcal{P}(\lambda)$ :

$$\Pi(s) = e^{-\lambda} \sum_{k=0}^{s} \frac{\lambda^k}{k!} \,.$$

| $s \setminus \lambda$ | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0.9048 | 0.8187 | 0.7408 | 0.6703 | 0.6065 | 0.5488 | 0.4966 | 0.4493 | 0.4066 |
| 1 | 0.9953 | 0.9825 | 0.9631 | 0.9384 | 0.9098 | 0.8781 | 0.8442 | 0.8088 | 0.7725 |
| 2 | 0.9998 | 0.9989 | 0.9964 | 0.9921 | 0.9856 | 0.9769 | 0.9659 | 0.9526 | 0.9371 |
| 3 | 1.0 | 0.9999 | 0.9997 | 0.9992 | 0.9982 | 0.9966 | 0.9942 | 0.9909 | 0.9865 |
| 4 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.9998 | 0.9996 | 0.9992 | 0.9986 | 0.9977 |
| 5 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.9998 | 0.9997 |
| 6 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |

| $s \setminus \lambda$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0.3679 | 0.1353 | 0.0498 | 0.0183 | 0.0067 | 0.0025 | 0.0009 | 0.0003 | 0.0001 |
| 1 | 0.7358 | 0.406 | 0.1991 | 0.0916 | 0.0404 | 0.0174 | 0.0073 | 0.003 | 0.0012 |
| 2 | 0.9197 | 0.6767 | 0.4232 | 0.2381 | 0.1247 | 0.062 | 0.0296 | 0.0138 | 0.0062 |
| 3 | 0.981 | 0.8571 | 0.6472 | 0.4335 | 0.265 | 0.1512 | 0.0818 | 0.0424 | 0.0212 |
| 4 | 0.9963 | 0.9473 | 0.8153 | 0.6288 | 0.4405 | 0.2851 | 0.173 | 0.0996 | 0.055 |
| 5 | 0.9994 | 0.9834 | 0.9161 | 0.7851 | 0.616 | 0.4457 | 0.3007 | 0.1912 | 0.1157 |
| 6 | 0.9999 | 0.9955 | 0.9665 | 0.8893 | 0.7622 | 0.6063 | 0.4497 | 0.3134 | 0.2068 |
| 7 | 1.0 | 0.9989 | 0.9881 | 0.9489 | 0.8666 | 0.744 | 0.5987 | 0.453 | 0.3239 |
| 8 | 1.0 | 0.9998 | 0.9962 | 0.9786 | 0.9319 | 0.8472 | 0.7291 | 0.5925 | 0.4557 |
| 9 | 1.0 | 1.0 | 0.9989 | 0.9919 | 0.9682 | 0.9161 | 0.8305 | 0.7166 | 0.5874 |
| 10 | 1.0 | 1.0 | 0.9997 | 0.9972 | 0.9863 | 0.9574 | 0.9015 | 0.8159 | 0.706 |
| 11 | 1.0 | 1.0 | 0.9999 | 0.9991 | 0.9945 | 0.9799 | 0.9467 | 0.8881 | 0.803 |
| 12 | 1.0 | 1.0 | 1.0 | 0.9997 | 0.998 | 0.9912 | 0.973 | 0.9362 | 0.8758 |
| 13 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.9993 | 0.9964 | 0.9872 | 0.9658 | 0.9261 |
| 14 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9998 | 0.9986 | 0.9943 | 0.9827 | 0.9585 |
| 15 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.9995 | 0.9976 | 0.9918 | 0.978 |
| 16 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9998 | 0.999 | 0.9963 | 0.9889 |
| 17 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.9996 | 0.9984 | 0.9947 |
| 18 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.9993 | 0.9976 |
| 19 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9997 | 0.9989 |
| 20 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9999 | 0.9996 |
| 21 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9998 |
| 22 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.9999 |
| 23 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |